



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 110233**

**TO: Jeffrey Parkin**  
**Location: cm1/8e15/8e12**  
**Art Unit: 1648**  
**Friday, December 12, 2003**

**Case Serial Number: 10/032162**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**CM1-6B02**  
**Phone: 305-9203**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Parkin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

THIS PAGE BLANK (US:70)

From: Parkin, Jeffrey  
Sent: Wednesday, December 10, 2003 5:44 PM  
To: STIC-Biotech/ChemLib  
Subject: U.S. Serial No. 10/032,162

RECEIVED  
DEC 11 2003

mej

Please search SEQ ID NOS.: 13, 15, and 17 from the aforementioned application (U.S. Serial No. 10/032,162) v. all relevant databases, including interference. Place results on BOTH disk and paper.  
Thanks!

JSP  
Au 1648  
CM01-8E15  
308-2227

13 AA 643  
15 | 579  
17 | 625

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 12/18/03  
Date Completed: 1/5/04  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 3  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: QSP  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THIS PAGE BLANK (US 710)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:24:08 ; Search time 38.2945 Seconds  
(without alignments)  
2665.163 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLWTVTYGVVPVKEATT.....ELDKWASLWNWFDTNNLWY 643

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq 19Jun03.\*

1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3462	100.0	643	22	AA1980
2	3295	95.2	625	22	AA1981
3	3091.5	89.3	850	16	AA1982
4	3082	89.0	883	22	AA1983
5	3069.5	88.7	855	18	AA1984
6	3069.5	88.7	855	20	AA1985
7	3056	88.3	579	22	AA1986
8	3041	87.8	847	21	AA1987
9	3036	87.7	842	24	AA1988

10	3026	87.4	842	23	AB06211	HIV Env isolate SF
11	2999	86.6	857	16	AA1989	gp120 from the HIV
12	2991.5	86.4	851	9	AA1990	HIV protein HT6.
13	2989.5	86.4	868	7	AA1991	Sequence of LAV vi
14	2986.5	86.3	868	23	AA1992	Lymphadenopathy-as
15	2985.5	86.2	868	7	AA1993	HIV virus env gene
16	2982.5	86.1	855	12	AA1994	HIV-1 BA-L clone-e
17	2982.5	86.1	855	19	AA1995	HIV-1 gp120 protei
18	2979	86.0	1188	17	AA1996	HIV-1 gp160-NYS p5
19	2978	86.0	856	18	AA1997	Human Immunodefici
20	2978	86.0	856	20	AA1998	Predicted amino ac
21	2978	86.0	856	22	AA1999	Amino acid sequenc
22	2978	86.0	863	14	AA2000	HTLV-III ENV-LOR 9
23	2977	86.0	854	21	AA2001	HIV-1 env protein.
24	2977	86.0	854	21	AA2002	HIV-1 coat protein
25	2977	86.0	854	22	AA2003	HIV gp41 DNA SEQ I
26	2977	86.0	854	22	AA2004	Protein encoded by
27	2977	86.0	854	23	AB073663	HIV-1 NL4-3 gp120
28	2974.5	85.9	856	16	AA2005	gp120 from the HIV
29	2974	85.9	856	7	AA2006	Sequence of envelo
30	2974	85.9	856	20	AA2007	HIV-1 env protein
31	2972.5	85.9	901	8	AA2008	Sequence encoded b
32	2972	85.8	855	19	AA2009	ENV protein contai
33	2972	85.8	855	21	AA2010	HIV-1 (ATCC CRL 85
34	2972	85.8	855	21	AA2011	HIV-1 (ATCC CRL 85
35	2972	85.8	855	24	AB057550	AIDS associated re
36	2972	85.8	855	24	AB057553	AIDS associated re
37	2972	85.8	856	7	AA2012	Sequence of the AI
38	2972	85.8	863	7	AA2013	Sequence of ARV-2
39	2972	85.8	863	13	AA2014	env gene decoded f
40	2970	85.8	856	13	AA2015	Modified HIV env g
41	2969	85.8	857	12	AA2016	HIV-1 (MN-STI) env
42	2969	85.8	857	18	AA2017	Human Immunodefici
43	2969	85.8	857	20	AA2018	Predicted amino ac
44	2968	85.7	863	7	AA2019	HTLV-III virus (HI
45	2967.5	85.7	860	18	AA2020	HIV-SF2 virus gp12

## ALIGNMENTS

RESULT 1  
AA1980  
ID AA1980 standard; Protein; 643 AA.  
XX  
AC AA1980;  
XX  
DT 05-APR-2001 (first entry)  
XX  
DE HIV-1 SOS gp140 glycoprotein.  
XX  
KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;  
KW gp140.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200100648-A1.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-US17267.  
XX  
PR 25-JUN-1999; 99US-0340992.  
XX  
PA (PROG-) PROGENICS PHARM INC.  
PA (AARO-) AARON DIAMOND AIDS RES CENT.  
XX  
PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;  
XX  
DR WPI, 2001-122993/13.  
DR N-PSDB; AAF28581.  
XX  
PT New viral envelope proteins, useful for producing vaccines to treat



Db 343 WNNTEGNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPPIRGQIRCSNITGLLLTRDG 402  
 Qy 421 GINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAVGIGAV 480  
 Db 403 GINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAVGIGAV 462  
 Qy 481 FLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNLLRAIEAQQRMQLTVMGIKQLQA 540  
 Db 463 FLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNLLRAIEAQQRMQLTVMGIKQLQA 522  
 Qy 541 RVLAVERYLGDQOLGIGCSGKLICCTAVPNWASWSNKS LDRINWNTMTWMEWEREIDNY 600  
 Db 523 RVLAVERYLGDQOLGIGCSGKLICCTAVPNWASWSNKS LDRINWNTMTWMEWEREIDNY 582  
 Qy 601 TSEIYTLIEESQNOQEKNEQELLELDKWSLWNNWFDITNWLWY 643  
 Db 583 TSEIYTLIEESQNOQEKNEQELLELDKWSLWNNWFDITNWLWY 625  
 ID AAR67724 standard; Protein; 850 AA.  
 AC AAR67724;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 07-SEP-1995 (first entry)  
 XX  
 DE gp120 from the HIV GNE8 isolate.  
 XX  
 KW HIV; human immunodeficiency virus; gp120; glycoprotein;  
 KW GNE clone; GNE8 isolate.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9428929-A1.  
 XX  
 PD 22-DEC-1994.  
 XX  
 PF 07-JUN-1994; 94WO-US06036.  
 XX  
 PR 07-JUN-1993; 93US-0072833.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Berman PW, Nakamura GR;  
 XX  
 WPI; 1995-036112/05.  
 N-PSDB; AAQ76018.  
 XX  
 PT Use of HIV gp 120 polypeptide(s) - for developing probes for the  
 PT analysis, prevention and therapy of HIV infection  
 XX  
 PS Claim 23; Page 27-30; 108pp; English.  
 XX  
 CC This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate  
 CC of HIV-1 (human immunodeficiency virus type 1). There are neutralising  
 CC epitopes in the V2 and C4 domains of gp120, in addition to the  
 CC neutralising epitopes in the V3 domain. Although the amino acid sequences  
 CC of the neutralising epitopes in the V2, V4 and C4 domains are variable,  
 CC the amount of variation is highly constrained. This facilitates the  
 CC design of HIV subunit vaccines that can induce antibodies that neutralise  
 CC the most common HIV strains for a given geographic region. This invention  
 CC provides a multivalent gp120 subunit vaccine where the gp120 present in  
 CC the vaccine is from at least two HIV isolates which have different amino  
 CC acid sequences for a neutralising epitope in these regions.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 850 AA;  
 Query Match 89.3%; Score 3091.5; DB 16; Length 850;  
 Best Local Similarity 89.7%; Pred. No. 2.6e-152;

Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;  
 Qy 2 EKLWTVYGVVWKEATITLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTE 61  
 Db 31 EKLWTVYGVVWKEATITLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTE 90  
 Qy 62 HFNWKNMVBQMOEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMBRG 120  
 Db 91 NFNWKNMVBQMOEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMBRG 150  
 Qy 121 EIKKCSNITTSIRDEVOKEYALFKLDVVDIDNNNTSYRLISCDTSVITQACPKISSEP 180  
 Db 151 EIKKCSNITTSIRDEVOKEYALFKLDVVDIDNNNTSYRLISCDTSVITQACPKISSEP 210  
 Qy 181 IPIHYCAPAGFAILLKCNKDTFNGKGPQCVSTVCTHGIRPVVSTQLLNLSLAEEVVI 240  
 Db 211 IPIHYCAPAGFAILLKCNKDTFNGKGPQCVSTVCTHGIRPVVSTQLLNLSLAEEVVI 270  
 Qy 241 RSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGRAFYTGTGIIIGDIRQAHCN 300  
 Db 271 RSANFSDNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGRAFYTGTGIIIGDIRQAHCN 330  
 Qy 301 SRAKNWDTLKOIVIKLREOPENKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSTOLPNST 360  
 Db 331 SSTKNWNTLKOIVIKLREHF-NKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSTOLPNST 389  
 Qy 361 WN-----NNTGSSNNTGNTITLPCRIKQIINMWQEVGKAMYAPPPIRGQIRCSNITGLL 415  
 Db 390 WNYTYTNWNTGSSNDT-GRNITLQCRINQIINMWQEVGKAMYAPPPIRGQIRCSNITGLL 448  
 Qy 416 LTRDGGINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAV 475  
 Db 449 LTRDGG-NNSETEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAV 507  
 Qy 476 GIGAVFLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNLLRAIEAQQRMQLTVMGI 535  
 Db 508 GIGAVFLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNLLRAIEAQQRMQLTVMGI 567  
 Qy 536 KOLQARVLAVERYLGDQOLGIGCSGKLICCTAVPNWASWSNKS LDRINWNTMTWMEWER 595  
 Db 568 KOLQARVLAVERYLGDQOLGIGCSGKLICCTAVPNWASWSNKS LDRINWNTMTWMEWER 627  
 Qy 596 EIDNYTSEIYTLIEESQNOQEKNEQELLELDKWSLWNNWFDITNWLWY 643  
 Db 628 EIDNYTSEIYTLIEESQNOQEKNEQELLELDKWSLWNNWFDITNWLWY 675  
 RESULT 4  
 AAB82761  
 ID AAB82761 standard; Protein; 883 AA.  
 AC AAB82761;  
 XX  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Ancestral HIV-1 group M, subtype B gp160 protein.  
 KW HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.  
 OS Human immunodeficiency virus type 1.  
 PN WO200160838-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 16-FEB-2001; 2001WO-US05288.  
 XX  
 PR 18-FEB-2000; 2000US-0183659.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Mullins JI, Rodrigo AG, Learn GH, Li F;  
 XX

DR MPI; 2001-536565/59.  
 XX N-PSDB; AAH26468.  
 PT Preparing an ancestral viral amino acid sequence useful as a vaccine  
 PT comprises determining a recent ancestor of a circulating virus by  
 PT maximum likelihood phylogeny analysis -  
 XX  
 PS Claim 8; Page 54; 89pp; English.  
 XX  
 CC The present sequence is that of an ancestral HIV-1 group M, subtype  
 CC B gp160 (env gene product) sequence. The invention provides  
 CC compositions and methods for determining ancestral viral gene  
 CC sequences and ancestral viral protein sequences for highly diverse  
 CC viruses, such as HIV-1. The methods use samples of circulating  
 CC viruses to determine an ancestral viral sequence by maximum  
 CC likelihood phylogeny analysis. In the present case, the ancestral  
 CC HIV-1 subtype B env sequence (see AAH26468) was determined using 38  
 CC subtype B sequences (obtained from 9 different countries) and 3  
 CC subtype D (outgroup) sequences. The distances between this ancestral  
 CC viral sequence and circulating strains used to determine it were on  
 CC average 12.3% (range: 8.0-21.0%) while the available specimens were  
 CC 17.3% different from each other (range: 13.3-23.2%). Thus, the  
 CC ancestor sequence was, on average, more closely related to any given  
 CC circulating virus than to any other variant. The ancestral gp160  
 CC sequence included a wide variety of immunogenically active peptides  
 CC when processed for antigen presentation; nearly all known subtype B  
 CC CTL epitope consensus amino acids were represented. Thus, an  
 CC immunogenic composition to this subtype B ancestor protein will  
 CC elicit broad neutralising antibody against HIV-1 isolates of the  
 CC same subtype, and will also elicit a broad cellular response  
 CC mediated by antigen-specific T-cells. A claimed vaccine composition  
 CC comprises a viral ancestor protein or its immunogenic fragment,  
 CC especially one derived from the HIV-1 group M subtype B gp160  
 CC ancestral protein.  
 XX  
 SQ Sequence 883 AA;

Query Match 89.0%; Score 3082; DB 22; Length 883;  
 Best Local Similarity 87.2%; Pred. No. 8.2e-152;  
 Matches 591; Conservative 19; Mismatches 32; Indels 36; Gaps 9;

QY 2 EKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTE 61  
 DB 1 EKLWTVYVYGVVWKEATTTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTE 90  
 QY 62 HFNWKNWYEQHEDIISLWDSQSLKPCVKLTPLCVTLNCKD---VNATNTWDS--- 114  
 DB 1 HFNWKNWYEQHEDIISLWDSQSLKPCVKLTPLCVTLNCKD---VNATNTWDS--- 150  
 QY 115 -----GTM--ERGEIKNCSENIITTSIRDEVOKEVALFYKLDVVPID-----NNNTSYRL 161  
 DB 151 TSSGGGFMGEKEGEIKNCSENIITTSIRDEVOKEVALFYKLDVVPIDNDNNNTWNTSYRL 210  
 QY 162 ISCDTSVITQACPKISPEPIPIHYCAPAGFAILKCNCKTENGKGPCKNVSTVQCTHGIRP 221  
 DB 211 INCNTSVITQACPKVSEPIPIHYCTPAGFAILKCNCKENGCTGCTNVSTVQCTHGIRP 270  
 QY 222 VVSTOLLNGLAEVEVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGR 281  
 DB 271 VVSTOLLNGLAEVEVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGR 330  
 QY 282 AFYTTGIIIGDIQAHNCISRAKWNNTLKOIVTKLREQFENK--TIVFNHSSGGDPEIVM 339  
 DB 331 ALVATGIIIGDIQAHNCISRAKWNNTLKOIVTKLREQFENK--TIVFNHSSGGDPEIVM 390  
 QY 340 HSFNCEGEFFYCNSTQLFNSTW-----NNTEGSNNT--EGNTITLPCRIKQIINMWQ 390  
 DB 391 HSFNCEGEFFYCNSTQLFNSTW-----NNTEGSNNT--EGNTITLPCRIKQIINMWQ 450  
 QY 391 EVGKAMYAPPIRGQIRCSNITGLLLTRDGINEN-----GTEIFRPGGDMRDNWRSEFY 446  
 DB 451 EVGKAMYAPPIRGQIRCSNITGLLLTRDGINEN-----GTEIFRPGGDMRDNWRSELY 510

447 KYKVKIEPLGVAPTCKKRRVVRKRAVG-ICAVFLGFLGAGSTMGASMTLTVOARL 505  
 DB 511 KYKVKIEPLGVAPTCKKRRVVRKRAVGMLGFLGAGSTMGASMTLTVOARQ 570  
 QY 506 LLSGIVQQNNLLRAIEAQRMLQLTVMGIKOLQARVLAVERYLGDQQLLGIWGCSSGLI 565  
 DB 571 LLSGIVQQNNLLRAIEAQRMLQLTVMGIKOLQARVLAVERYLKDQQLLGIWGCSSGLI 630  
 QY 566 CTAVPWNASWNSKSLDRINNNMTWMEWEREDNTYSETYLTIEESQOQKNEQELLEL 625  
 DB 631 CTAVPWNASWNSKSLDRINNNMTWMEWEREDNTYGLTYLTIEESQOQKNEQELLEL 690  
 QY 626 DKWASLWNNWFDITNWLWY 643  
 DB 691 DKWASLWNNWFDITNWLWY 708

RESULT 5  
 AAW11581  
 ID AAW11581 standard; Protein; 855 AA.  
 XX AAW11581;  
 AC AAW11581;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 25-MAR-1997 (first entry)  
 DT  
 XX Human Immunodeficiency Virus-1 strain BA-L envelope protein.  
 DE  
 XX Acquired immune deficiency syndrome; AIDS; envelope protein;  
 KW env gene; vaccine.  
 KW  
 XX Human immunodeficiency virus type 1 (strain BA-L).  
 OS  
 XX US5576000-A.  
 PN  
 XX 19-NOV-1996.  
 PD  
 XX 15-FEB-1995; 95US-0388809.  
 PP  
 XX 17-OCT-1990; 90US-0599491.  
 PR 25-FEB-1993; 93US-0022835.  
 PR 15-FEB-1995; 95US-0388809.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;  
 PI Popovic M, Reitz MS;  
 PI  
 XX WPI; 1997-011206/01.  
 DR N-PSDB; AAT58551.  
 DR  
 XX New isolated envelope protein of HIV-1 strain BA-L and recombinant  
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,  
 PT typical of US clinical isolates  
 PT  
 XX Claim 1; Fig 9; 86pp; English.  
 PS  
 XX A HindIII fragment of unintegrated viral DNA representing the BA-L  
 CC genome was cloned into lambda phage Charon 28 DNA from total DNA of  
 CC peripheral blood lymphocytes infected with and producing HIV-1(BA-L).  
 CC A positive clone was selected by hybridisation using a HIV-1 env  
 CC probe. This clone, designated BA-Li, was found to contain the  
 CC entire gene for the envelope protein on a 2.8 kb HindIII-XbaI  
 CC fragment and a 0.4 kb EcoRI-HindIII fragment. When cloned together  
 CC these fragments comprise the env gene, as well as the coding regions  
 CC for rev and the rev-responsive element of env, both necessary for  
 CC efficient expression in eukaryotic cells. The claimed recombinantly  
 CC produced envelope protein can be used as an immunogen for raising  
 CC antibodies against HIV.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX Sequence 855 AA;

XX	17-OCT-1990;	90US-0599491.
PR	25-FEB-1993;	93US-0022835.
PR	15-FEB-1995;	95US-0388809.
PR	14-MAY-1996;	96US-0647714.
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
PA	Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD; Popovic M, Reitz MS;	
XX	WPI; 1999-152779/13.	
DR	N-PSDB; AAX04767.	
XX	DNA encoding env protein of the human immune deficiency virus isolate BA-L - useful for producing protein for use in vaccines, as assay reagent and to generate antibodies	
XX	Example 1; Fig 9A-C; 87pp; English.	
PS	The present sequence represents the envelope protein of the BA-L (ATCC 40890) strain of Human immunodeficiency virus type 1 (HIV-1) strain MN-STL. BA-L is more typical of United States isolates of HIV-1 than previously known strains. Recombinant, complete env protein of the BA-L strain is used as a vaccine component and for immunotherapy of existing HIV infections, to detect HIV-specific antibodies, e.g. in donated blood, and as an immunogen to raise specific antibodies, for HIV-1 diagnosis.	
XX	Sequence 855 AA;	
SQ	Query Match 88.7%; Score 3069.5; DB 20; Length 855; Best Local Similarity 88.1%; Pred. No. 3.5e-151; Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;	
QY	2 EKLWTYYGVVWKAEATTTLCASDAKAYDTEVHVNWATHACVTPDPNPQEVLENVTE 61	
DB	31 EKLWTYYGVVWKAEATTTLCASDKAYDTEVHVNWATHACVTPDPNPQEVLEKNVTE 90	
QY	62 HENMKNNNVEOMOEIIISLDOSLKPCVKLPLCVTLNCKDV-----NATNTTNDSEG 115	
DB	91 NFNMKNNNVEQMHEDIISLDOSLKPCVKLPLCVTLNCTLRNATNGNDTNTSSRG 150	
QY	116 TWERGEIKNCSEFNITTSIRDEQKYALFYKLDVVPIDNN-NTSYRLISCDTSVITOACP 174	
DB	151 MYGGEWMKCSEFNITTNIRGKVQEYALFYKLDIADPDDNSNNRYLISCNTSVITOACP 210	
QY	175 KISFEPIPIHYCAPAGFAILKCNKDFTFNGKGPKCNVSTVQCTHGIRPVVSTQLLLNGSLA 234	
DB	211 KVSFEPIPIHYCAPAGFAILKCKDKFKNGKPGCTNVSTVQCTHGIRPVVSTQLLLNGSLA 270	
QY	235 EEDEVIRSDFNTNAKTIIVOLKESVEINCTRPNNTRKSIIHIGPGRAFYTTCGIIIDIR 294	
DB	271 EEDEVIRSFANADNAKIIVQLNESVEINCTRPNNTRKSIIHIGPGRAFYTTCGIIIDIR 330	
QY	295 QAHCNISRAKWDTLKOIVIKLREOFENKTIVFNHSSGGDPEIVMHSFNCEGFEFFVCNST 354	
DB	331 QAHCNLISRAKWNLTNKVIIVIKLREQGNKTIIVFKUSSGGDDELVTSHSFCNGGFEFFVCNST 390	
QY	355 QLFNSTWNNTSGSNNTGNTITLPCRIFOIINNMQEVGKAMYAPPIRGIQRCSSNITGL 414	
DB	391 QLFNSTWNVTESNNTVENNTITLPCRIFQIINNMQEVGVMYAPPIRGIQRCSSNITGL 450	
QY	415 LLTROGGINENGTEIFRPGGGDMRDNRWSEFYKYKKVKEPLGVAPTCKGRVVQREKRA 474	
DB	451 LLTROGGPEDNTEVFRPGGGDMRDNRSELYKYKKVKEPLGVAPTCKARRVVQREKRA 510	
QY	475 VGGIAVFELFGAAGSTMGAASMTLTVOARLLLSGTVQQOONLLRAIEAQORMQLQTVWG 534	
DB	511 VGGIAVFELFGAAGSTMGAAAWTLTVQARLLLSGTVQQOONLLRAIEAQOHLLQTVWG 570	
QY	535 IKQLOARVLAVERYLGDOQLLIGWCSGKLI CCTAVPMNASMSNKSGLDRIWNNTMWEE 594	

Db 571 IKOLARVLAVRYLRDQQLLGIWCGSKLICCTTAVPWNASWNSKLNKINWNTMTIEWD 630

QY 595 REIDNVTSEIYTLIESQOKEQNEQLLELDKWSLNNWFDITNLWY 643

Db 631 REINNTTSIIYSLIESQOKEQNEQLLELDKWSLNNWFDITNLWY 679

RESULT 7

AA61506

ID AAB61506 standard; Protein; 579 AA.

XX AAB61506;

AC

XX

DT 05-APR-2001 (first entry)

XX

DE HIV-1 deltaV12\* SOS gp140 glycoprotein.

XX

KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41; gp140.

KW

Human immunodeficiency virus type 1.

XX

PN WO200100648-A1.

XX

XX 04-JAN-2001.

PD

XX 23-JUN-2000; 2000WO-US17267.

PF

XX 25-JUN-1999; 99US-0340992.

PR

XX (PROG-) PROGENICS PHARM INC.

PA (AARO-) AARON DIAMOND AIDS RES CENT.

PA

XX Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;

PI

XX WPI; 2001-122993/13.

DR

DR N-PSDB; AAF28582.

DR

XX New viral envelope proteins, useful for producing vaccines to treat

PT human immunodeficiency virus-1 infections, comprises amino acid

PT sequence mutations such that viral transmembrane-surface protein

PT complex is more stable

PT

XX Disclosure; Fig 14; 109pp; English.

PS

XX The present invention relates to a viral envelope protein. The viral

XX envelope protein comprises a viral surface protein (e.g. glycoprotein

CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in

CC which the viral envelope protein contains one or more amino acid sequence

CC mutations that enhance the stability of the complex formed between the

CC viral surface and transmembrane proteins. The viral envelope protein can

CC be used in the treatment of viral infection e.g. HIV-1 infection. The

CC present sequence is HIV-1 deltaV12\* SOS gp140, which was used in the

CC present invention.

XX

SQ Sequence 579 AA;

Query Match 88.3%; Score 3056; DB 22; Length 579;

Best Local Similarity 89.3%; Pred. No. 1.3e-150;

Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;

QY 1 VEKLWTVYGVVPWKKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEWLENTV 60

Db 1 VEKLWTVYGVVPWKKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEWLENTV 60

QY 61 EHFNMWNNVQEQEDIIISLWQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGMERG 120

Db 61 EHFNMWNNVQEQEDIIISLWQSLKPCVKLTPLCGA----- 98

QY 121 EIKNCSFNITTSRDEVOKEVALFYKLDVVPIDNNNTSYELISCDTSVITQACPKISFEP 180

Db 99 -----GCDTSVITQACPKISFEP 116

QY 181 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 240

Db 117 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 176

QY 241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGPGRAFYTTGEIIGDIRQAHCVI 300

Db 177 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGPGRAFYTTGEIIGDIRQAHCVI 236

QY 301 SRAKWNDTLKQIVIKLREQFENKTIIVFNHSSGGDPEIIVMHSFNCGEFFYCNSTOLFNST 360

Db 237 SRAKWNDTLKQIVIKLREQFENKTIIVFNHSSGGDPEIIVMHSFNCGEFFYCNSTOLFNST 296

QY 361 MNNTEGSNNTGNTITLPCRIKQIINMMQEVGKAMYAPPPIRGQIRCSSNITGLLTRDG 420

Db 297 MNNTEGSNNTGNTITLPCRIKQIINMMQEVGKAMYAPPPIRGQIRCSSNITGLLTRDG 356

QY 421 GINENGTEIFRPGGDMRDNRSEFYKYKVVKIEPLGVAPTCKRVRVQREKRAVGIGAV 480

Db 357 GINENGTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTCKRVRVQREKRAVGIGAV 416

QY 481 FLGFLGAAGSTMGAASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQMLQLTVMGIKQLOA 540

Db 417 FLGFLGAAGSTMGAASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQMLQLTVMGIKQLOA 476

QY 541 RVLAVERYLGDQQLLGIWCGSKLICCTTAVPWNASWNSKSLDRIMNNMTMWEEREIDNY 600

Db 477 RVLAVERYLGDQQLLGIWCGSKLICCTTAVPWNASWNSKSLDRIMNNMTMWEEREIDNY 536

QY 601 TSEIYTLIESQOKEQNEQLLELDKWSLNNWFDITNLWY 643

Db 537 TSEIYTLIESQOKEQNEQLLELDKWSLNNWFDITNLWY 579

RESULT 8

AA97073

ID AAY97073 standard; protein; 847 AA.

XX

AC AAY97073;

XX

DT 31-OCT-2000 (first entry)

XX

DE Variant HIV-1 SF162 Env gp160.

XX

KW HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;

KW CD4 binding region; V1/V2 loop; bridging sheet.

XX

OS Human immunodeficiency virus type 1 isolate SF162.

XX WO200039303-A2.

XX

PD 06-JUL-2000.

XX

PF 30-DEC-1999; 99WO-US31272.

XX

PR 31-DEC-1998; 98US-0114495.

PR 29-SEP-1999; 99US-0156670.

XX

XX (CHIR ) CHIRON CORP.

PA Barnett S, Hartog K, Martin E;

XX WPI; 2000-465745/40.

DR

XX Novel modified HIV Env polypeptides useful as immunizing agents and for

PT preparing a vaccine to elicit an immune response against a broad range

PT of HIV subtypes

PS

XX Claim 5; Page 115-117; 139pp; English.

XX

CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope

CC (Env) polypeptides having an amino acid deleted or replaced in the region

CC corresponding to residues 420-436 or 119-123 and 199-210 relative to

CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV

CC strain SP162, with numbering relative to isolate HXB-2. The Env  
CC polypeptides are modified so as to expose at least part of the CD4  
CC binding region. The modified HIV Env polypeptides, coding polynucleotides  
CC and constructs, further comprising an adjuvant, are used for inducing an  
CC immune response in an individual. The method involves administering a  
CC first composition comprising a polynucleotide encoding the Env  
CC polypeptide in a priming step and administering a second composition  
CC comprising a modified Env polypeptide as a booster in an amount  
CC sufficient to induce an immune response in the individual. The first  
CC and/or second composition further comprises an adjuvant (claimed). The  
CC intracellularly produced Env polypeptides can be used for a number of  
CC diagnostic and therapeutic purposes to determine the presence of reactive  
CC antibodies and/or Env proteins in a biological sample to aid in the  
CC diagnosis of HIV infection or disease status or as measure of response to  
CC immunization.

XX SQ Sequence 847 AA;

Query Match 87.8%; Score 3041; DB 21; Length 847;  
Best Local Similarity 88.7%; Pred. No. 1e-149;  
Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

QY 1 VEKLVVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 60  
DB 30 VEKLVVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 89  
QY 61 EHFNNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMR 119  
DB 90 EHFNNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMR 149  
QY 120 GEIKNCSFNITTSIRDEQVEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKNIFE 179  
DB 150 GEIKNCSFKVTTISIRNMQEYALFYKLDVVPIDNDNTSYKLNCNTSVITQACPKNIFE 209  
QY 180 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEV 239  
DB 210 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEV 269  
QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGEEIIGDIRQACHN 299  
DB 270 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGEEIIGDIRQACHN 329  
QY 300 ISRAKNDTLKQIVIKLREQENKTIIVNHSAGDPEIVMHSFNCBGEFFYCNSTQLFNS 359  
DB 330 ISGKKNWNLKQIVIKLREQENKTIIVNHSAGDPEIVMHSFNCBGEFFYCNSTQLFNS 389  
QY 360 TWNNTEGNNTEGNTITLPCRIKQIINWQEVGKAMYAPPIRGOIRCSSNITGLLLTRD 419  
DB 390 TW-NNTIGPNNTG-TITLPCRIKQIINWQEVGKAMYAPPIRGOIRCSSNITGLLLTRD 447  
QY 420 GGIN-ENGTEIFRPGGDMRDNWRSEFYKVKVIEPLGVAPTCKRRVQREKRAVGIG 478  
DB 448 GGEISNTTEIFRPGGDMRDNWRSELYKVKVIEPLGVAPTCKRRVQREKRAVTLG 507  
QY 479 AVFLGFLGAAGSTGMAASMTITVQARLLSGIVQQQNNLLRAIEAQQRLQITVWGIKQL 538  
DB 508 AMFLGFLGAAGSTGMAASMTITVQARLLSGIVQQQNNLLRAIEAQQRLQITVWGIKQL 567  
QY 539 QARVLAVERYLGDQQLIGWCSGKLICTTAVPNWASNSKSLDRINWNTWMEWERID 598  
DB 568 QARVLAVERYLGDQQLIGWCSGKLICTTAVPNWASNSKSLDRINWNTWMEWERID 627  
QY 599 NYTSEIYTLIESQNOQEKNEQELLELDKASLWNPDIINWLY 643  
DB 628 NYTNLIYTLIESQNOQEKNEQELLELDKASLWNPDIINWLY 672

RESULT 9  
ID ABU66565  
XX ABU66565 standard; Protein; 842 AA.  
AC ABU66565;  
XX

DT 22-MAY-2003 (first entry)  
XX Human immunodeficiency virus (HIV) envelope (env) protein #1.  
DE Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;  
XX vaccine; gene therapy; packaging cell line; humoral immune response;  
KW cellular immune response; gene delivery vector; DNA immunisation;  
KW envelope protein; env.  
XX Human immunodeficiency virus.  
XX WO2003004657-A1.  
XX 16-JAN-2003.  
XX 05-JUL-2002; 2002WO-US21421.  
XX 05-JUL-2001; 2001US-303192P.  
PR 31-AUG-2001; 2001US-316860P.  
PR 16-JAN-2002; 2002US-349728P.  
PR 16-JAN-2002; 2002US-349793P.  
PR 16-JAN-2002; 2002US-349871P.  
XX (CHIR ) CHIRON CORP.  
PA Zur Megede J, Barnett SW, Lian Y;  
XX WPI; 2003-221602/21.  
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type  
PT C polypeptides, useful as immunogenic compositions or vaccines for  
PT generating humoral or cellular immune responses against HIV in a  
PT subject, especially humans -  
XX Disclosure; Fig 2A-C; 262pp; English.  
PS The invention describes a synthetic polynucleotide encoding 2 or more  
XX immunogenic HIV polypeptides, where at least 2 of the polypeptides are  
CC derived from different HIV subtypes. The polynucleotide is useful for  
CC immunisation, generation of packaging cell lines, or production of HIV  
CC polypeptides. The polynucleotide and its encoded proteins are useful as  
CC immunogenic compositions or vaccines for generating humoral or cellular  
CC immune responses against HIV in a subject, or for inducing neutralising  
CC antibodies against HIV. The gene delivery vector comprising the  
CC polynucleotide is also useful for DNA immunisation of, or for  
CC generating an immune response (e.g. a humoral or cellular immune  
CC response) in, a subject such as a mammal, particularly a human. This  
CC is the amino acid sequence of a human immunodeficiency virus (HIV)  
CC envelope (env) protein.  
XX SQ Sequence 842 AA;  
Query Match 87.7%; Score 3036; DB 24; Length 842;  
Best Local Similarity 88.5%; Pred. No. 1.9e-149;  
Matches 571; Conservative 29; Mismatches 41; Indels 4; Gaps 4;  
QY 1 VEKLVVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 60  
DB 25 VEKLVVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 84  
QY 61 EHFNNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMR 119  
DB 85 EHFNNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMR 144  
QY 120 GEIKNCSFNITTSIRDEQVEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKNIFE 179  
DB 145 GEIKNCSFKVTTISIRNMQEYALFYKLDVVPIDNDNTSYKLNCNTSVITQACPKNIFE 204  
QY 180 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEV 239  
DB 205 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEV 264  
QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGEEIIGDIRQACHN 299

Db 265 IRSEFTDNTAKTIIIVOLKESVEINCTRPNNTRKSTITIGPRAFYATGDIIGDIRQAHCN 324  
 Qy 300 ISAKWNTDLKQIVIKREOFENKTIIVFNHSSGGDPEIIVHNSFCNCEGFEFFYCNSTOLFNS 359  
 Db 325 ISGEKWNNTLKQIVTKLQAFGNKTIIVFKQSGGDPPEIIVHNSFCNCEGFEFFYCNSTOLFNS 384  
 Qy 360 TNNNTGSGNNTGNTITLPCRIKQIINNMQEYGVKAMYPPIRGQIRCSNITGLLTLTRD 419  
 Db 385 TW-NTTIGPNTNG-TITLPCRIKQIINNMQEYGVKAMYPPIRGQIRCSNITGLLTLTRD 442  
 Qy 420 GGIN-ENGTEIFRPGGDMRDNRSFYKYVVKIEPLGVAPTCKRRVVQREKRAVGIG 478  
 Db 443 GKEISNTTIEIFRPGGDMRDNRSFYKYVVKIEPLGVAPTCKRRVVQREKRAVTLG 502  
 Qy 479 AVFLGFLGAAGSTMGASMTLTIVQARLLLSGIVQOQNLLRAEAQOHLQLTVMGIKQL 538  
 Db 503 AMFLGFLGAAGSTMGARSLLTIVQARQLLSGIVQOQNLLRAEAQOHLQLTVMGIKQL 562  
 539 QARVLAVERYLGDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDRIWNNNTWMEWEREID 598  
 563 QARVLAVERYLGDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDRIWNNNTWMEWEREID 622  
 599 NYTSEIYTLIEESQOQKNEQELLELDKWSLWNNWFDITNMLWY 643  
 623 NYTNLTIVTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWY 667

RESULT 10

ABB06211  
 ID ABB06211 standard; Protein; 842 AA.

AC ABB06211;

DT 15-MAY-2002 (first entry)

DE HIV Env isolate SFI62 amino acid sequence.

KW Human immunodeficiency virus type C; antigenic HIV type C protein;

KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;

KW immunostimulant; gene therapy.

OS Human immunodeficiency virus type C.

OS Synthetic.

PN W0200204493-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.

XX 05-JUL-2000; 2000US-0610313.

XX (CHIR ) CHIRON CORP.

XX (UYST-) UNIV STELLENBOSCH.

XX Zur Megede J, Barnett SW, Engelbrecht S, Van Renaburg EJ;

XX WPI; 2002-154920/20.

XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful

XX in applications including DNA immunization or generation of packaging

XX cell lines, particularly in gene therapy

XX Disclosure; Fig 105; 233pp; English.

XX The present invention describes expression cassettes comprising a

XX polynucleotide sequence encoding a polypeptide comprising immunogenic

XX HIV type C polypeptides. The expression cassettes comprise any of the

XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or

XX Nef (i). (i) have immunostimulant activity and can be used in gene

XX therapy. The HIV type C polynucleotides are useful in applications

XX including DNA immunisation, generation of packaging cell lines, and

CC production of HIV Type C proteins. The polynucleotides are particularly  
 CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
 CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 842 AA;

Query Match 87.4%; Score 3026; DB 23; Length 842;  
 Best Local Similarity 88.2%; Pred. No. 6.2e-149;  
 Matches 569; Conservative 29; Mismatches 43; Indels 4; Gaps 4;

Qy 1 VEKLVVTVYVYGVPMKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVT 60

Db 25 VEKLVVTVYVYGVPMKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVT 84

Qy 61 EHPNKNMNMVQMDEDIISLMDOSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMER 119

Db 85 ENFNMMNNMNMVQMDEDIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTMER 144

Qy 120 GEIKKCSFNITTSIRDEVOKEYALFYKLDVDPIDNNNTSYRLISCDTSVITQACPISPE 179

Db 145 GEIKKCSFKVTTTSIRNKQKEYALFYKLDVDPIDNDNTSYKLINCNTSVTTQACPVSPE 204

Qy 180 PPIHYCAPAGPAILKCNDDKTFNGKPCXKXVSTVQCTHGIRPVVSTQLLNGSLABEEVV 239

Db 205 PPIHYCAPAGPAILKCNDDKTFNGKPCXKXVSTVQCTHGIRPVVSTQLLNGSLABEEVV 264

Qy 240 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSHIIGPRAFYTTGIIIGDIRQAHCN 299

Db 265 IRSEFTDNTAKTIIVQLKESVEINCTRPNNTRKSTTIGPRAFYATGDIIGDIRQAHCN 324

Qy 300 ISRAKWNNTLKOIVIKREOFENKTIIVFNHSSGGDPEIIVHNSFCNCEGFEFFYCNSTOLFNS 359

Db 325 ISGEKWNNTLKQIVTKLQAFGNKTIIVFKQSGGDPPEIIVHNSFCNCEGFEFFYCNSTOLFNS 384

Qy 360 TNNNTGSGNNTGNTITLPCRIKQIINNMQEYGVKAMYPPIRGQIRCSNITGLLTLTRD 419

Db 385 TW-NTTIGPNTNG-TITLPCRIKQIINNMQEYGVKAMYPPIRGQIRCSNITGLLTLTRD 442

Qy 420 GGIN-ENGTEIFRPGGDMRDNRSFYKYVVKIEPLGVAPTCKRRVVQREKRAVGIG 478

Db 443 GKEISNTTIEIFRPGGDMRDNRSFYKYVVKIEPLGVAPTCKRRVVQREKRAVTLG 502

Qy 479 AVFLGFLGAAGSTMGASMTLTIVQARLLLSGIVQOQNLLRAEAQOHLQLTVMGIKQL 538

Db 503 AMFLGFLGAAGSTMGARSLLTIVQARQLLSGIVQOQNLLRAEAQOHLQLTVMGIKQL 562

Qy 539 QARVLAVERYLGDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDRIWNNNTWMEWEREID 598

Db 563 QARVLAVERYLGDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDRIWNNNTWMEWEREID 622

Qy 599 NYTSEIYTLIEESQOQKNEQELLELDKWSLWNNWFDITNMLWY 643

Db 623 NYTNLTIVTLIEESQOQKNEQELLELDKWSLWNNWFDISKWLWY 667

RESULT 11

AAR67725

ID AAR67725 standard; Protein; 857 AA.

XX AAR67725;

XX 25-MAR-2003 (updated)

DT 11-SEP-1995 (first entry)

XX gp120 from the HIV GNE16 isolate, clone 1.

XX HIV; human immunodeficiency virus; gp120; glycoprotein;

XX GNE clone; GNE16 isolate.

XX Human immunodeficiency virus type 1.

XX WO9428929-A1.





180 PIPHYCAPAGFAILLKNDXTFNGKGPCKXNVSTVQCTHGIRPVVSTQLLNGSLABEEVV 239  
 212 PIPHYCAPAGFAILLKCNKTFNGTCTVSTVQCTHGIRPVVSTQLLNGSLABEEVV 271  
 240 IRSDFNFTNAKTIIVQLKESVEINCTRPNNNTKSHIGPGRAPYTTGIIIGDIRAHGN 299  
 272 IRSENFDTNVKTIIVQLNASVQINCTRPNNNTKSHIGPGRAPYTTGIIIGDIRAHGN 331  
 300 ISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDDPEIIVHSEFNCSEGEPEYCNSTOLFNS 359  
 332 LSSAQWNTLQVVKLREOFDNKTIIVFTSSSGDDPEIIVHSEFNCSEGEPEYCNSTOLFNS 391  
 360 TNNNTSGSNNTBGN-TITLPCRIKQIINNMQEYGRAMYAPPPIRGQIRCSSNITGLLLTR 418  
 392 TW-NSTEGSNNTGNDTITLPCRIKQIIVNMVQEVGRAMYAPPISGQIKCSNITGLLLTR 450  
 419 DGGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVVQREKRAVGI 477  
 451 DGGEDTNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVVQREKRAVGI 510  
 478 GAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQQNNLLRAIEAQOQMLQITVNGIKQ 537  
 511 GALFLGFLGAGSTMGASMTLTVOARLLLSGIVQQNNLLRAIEAQOQMLQITVNGIKQ 570  
 538 LQARLAVERYLGDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDRINNMTWMEWREI 597  
 571 LQARLAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDRINNMTWMEWREI 630  
 598 DNTSEIYTLIESQOQKNEQELLELDKWSLWNNFDTNNLWY 643  
 631 NNTSLIHSIESQOQKNEQELLELDKWSLWNNFDTNNLWY 676

RESULT 13  
 AAP60422  
 ID AAP60422 standard; Protein; 868 AA.  
 XX  
 AC AAP60422;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 20-AUG-1991 (first entry)  
 XX  
 DE Sequence of LAV virus env protein.  
 XX  
 KW AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.  
 XX  
 OS Lymphadenopathy virus.  
 XX  
 MO8602383-A.  
 XX  
 PD 24-APR-1986.  
 XX  
 PF 18-OCT-1985; 85WO-EP00548.  
 XX  
 PR 18-OCT-1984; 84GB-0016013.  
 PR 16-NOV-1984; 84GB-0029099.  
 PR 21-JAN-1985; 85GB-0001473.  
 XX  
 (INSP ) INST PASTEUR.  
 PA (CNRS ) CENT NAT RECH SCIENTIF.  
 XX  
 Montagnier L, Krust B, Charnaret F, Chermann JC, Barresinou F;  
 PI Alizon M, Sonigo P;  
 XX  
 DR WPI; 1986-119166/18.  
 DR N-PSDB; AAN60365.  
 XX  
 Purified glyco:protein and peptide(s) - are recognised by sera contg.  
 PT antibodies against lymphadenopathy virus and useful in detecting  
 PT AIDS antibodies or in vaccines  
 XX  
 Disclosure; Fig 4; 75pp; English.  
 PS  
 PS  
 XX

CC The inventors claim a polypeptide which is recognised by sera of  
 CC human origin contg. antibodies against the virus of  
 CC lymphadenopathies (LAV) or acquired immune deficiency syndrome  
 CC (AIDS). Also claimed are various peptides corresp. to the AA  
 CC sequences deducible from proteins encoded by LAV DNA, defined by  
 CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance  
 CC with a formula given in the specification.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 868 AA;  
 Query Match 86.4%; Score 2989.5; DB 7; Length 868;  
 Best Local Similarity 86.6%; Pred. No. 4.9e-147;  
 Matches 568; Conservative 33; Mismatches 40; Indels 15; Gaps 8;  
 QY 2 EKLWTVVYGVVPMKREATTTLFCASDAKAYDEVHNVATHACVPTDPNPQEVVLNVTE 61  
 DB 39 EKLWTVVYGVVPMKREATTTLFCASDAKAYDEVHNVATHACVPTDPNPQEVVLNVTE 98  
 QY 62 HFNMMKNNVVEOMQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT----TNDSG- 115  
 DB 99 NFNMMKNDVVEOMQEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTSSNTSSGE 158  
 QY 116 -TWERGEIKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 174  
 DB 159 MMEKEGEIKNCSEFNITTSIRGVQKEYAFYKLDIIPIDNTTSTYLTSCNTSVITQACP 218  
 QY 175 KISFEPPIHYCAPAGFAILLKNDKTFNGKGPCKXNVSTVQCTHGIRPVVSTQLLNGSLA 234  
 DB 219 KVSPEPIHYCAPAGFAILLKCNKTFNGTCTVSTVQCTHGIRPVVSTQLLNGSLA 278  
 QY 235 EEEVIRSDNFTNAKTIIVQLKESVEINCTRPNNNTKSHI--GPGRAFYTGTGIIIGD 292  
 DB 279 EEEVIRSANFTDNKTIIVQLNQSVEINCTRPNNNTKSHIRIQPGRAFVTIGK-IGN 337  
 QY 293 IQAHCHNISRAKWNATLKOIVIKLREOF-ENKTIIVFNHSSGGDPEIIVHSEFNCSEGEFFYC 351  
 DB 338 MQAHCHNISRAKWNATLKOIVIKLREOFENKTIIVFNHSSGGDPEIIVHSEFNCSEGEFFYC 397  
 QY 352 NSTQLFNSNTWNN---NTEGSNNTG-NTITLPCRIKQIINNMQEYGRAMYAPPPIRGQIR 407  
 DB 398 NSTQLFNSNTWNSWTSGSNTGSDTITLPCRIKQIINNMQEYGRAMYAPPISQIR 457  
 QY 408 SSNTGLLITRDGGINENCTEIRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRV 467  
 DB 458 SSNTGLLITRDGGNNNGSEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCKRRV 517  
 QY 468 VOREKRAVGIGAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQQNNLLRAIEAQOQ 527  
 DB 518 VOREKRAVGIGALFLGFLGAGSTMGARSMTLTVOARLLLSGIVQQNNLLRAIEAQOHL 577  
 QY 528 LQITVWGIKQLOARVLAVERYLGDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDRINN 587  
 DB 578 LQITVWGIKQLOARILAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDRINN 637  
 QY 588 MTWMEWEREIDNVTSEIYTLIESQOQKNEQELLELDKWSLWNNFDTNNLWY 643  
 DB 638 MTWMEWDROINNTYSLIHSIESQOQKNEQELLELDKWSLWNNFDTNNLWY 693  
 RESULT 14  
 AAOL9389  
 ID AAOL9389 standard; Protein; 868 AA.  
 XX  
 AC AAOL9389;  
 XX  
 DT 10-DEC-2002 (first entry)  
 XX  
 DE Lymphadenopathy-associated virus env protein.  
 XX  
 KW Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 12, 2003, 12:28:04 ; Search time 14.6215 Seconds  
(without alignments)  
1660.670 Million cell updates/sec

Title: US-10-032-162-13

Percent score: 3462  
Sequence: 1 VEKLMVTYGVGPVWKEATT.....ELDKRASIWNPDITNIMLWY 643Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

1 number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiletest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3091.5	89.3	850	2	US-08-448-603A-28
2	3091.5	89.3	850	3	US-09-134-075-28
3	3091.5	89.3	850	4	US-09-492-739-28
4	3069.5	88.7	855	1	US-08-022-835-6
5	3069.5	88.7	855	1	US-08-388-809-6
6	3069.5	88.7	855	2	US-08-647-714-6
7	3027.5	87.4	855	2	US-07-956-483-14
8	2999	86.6	857	2	US-08-448-603A-30
9	2999	86.6	857	4	US-09-134-075-30
10	2999	86.6	857	4	US-09-492-739-30
11	2992.5	86.4	861	1	US-08-127-499A-14
12	2992.5	86.4	861	1	US-08-482-847-14
13	2992.5	86.4	861	3	US-07-956-483-10
14	2992.5	86.4	861	3	US-08-472-240A-7
15	2992.5	86.4	861	3	US-08-472-240A-7
16	2992.5	86.4	861	4	US-08-817-441-103
17	2984.5	86.2	861	3	US-07-956-483-16
18	2978	86.0	856	1	US-08-022-835-2
19	2978	86.0	856	1	US-08-388-809-2
20	2978	86.0	856	2	US-08-647-714-2
21	2978	86.0	856	3	US-07-956-483-11
22	2978	86.0	856	3	US-09-124-900-9
23	2978	86.0	863	3	US-08-463-210-11
24	2977	86.0	826	1	US-08-375-510-2
25	2977	86.0	826	2	US-08-487-651-2
26	2977	86.0	854	4	US-09-309-572-23
27	2969	85.8	857	1	US-08-022-835-4

28	2969	85.8	857	1	US-08-388-809-4	Sequence 4, Appli
29	2969	85.8	857	2	US-08-647-714-4	Sequence 4, Appli
30	2966	85.7	887	3	US-08-472-240A-5	Sequence 5, Appli
31	2964	85.6	855	3	US-07-956-483-15	Sequence 15, Appli
32	2964	85.6	856	2	US-07-916-098A-2	Sequence 2, Appli
33	2963.5	85.6	839	2	US-08-472-240A-10	Sequence 10, Appli
34	2953	85.3	856	4	US-09-337-387-11	Sequence 11, Appli
35	2947.5	85.1	880	2	US-08-788-815-7	Sequence 7, Appli
36	2947.5	85.1	880	4	US-09-157-963-7	Sequence 11, Appli
37	2935	84.8	856	1	US-08-375-100-1	Sequence 1, Appli
38	2923	84.4	887	3	US-08-472-240A-6	Sequence 6, Appli
39	2917	84.3	865	3	US-07-956-483-13	Sequence 13, Appli
40	2910.5	84.1	726	4	US-09-337-387-3	Sequence 3, Appli
41	2907	84.0	887	3	US-08-472-240A-2	Sequence 2, Appli
42	2897.5	83.7	759	4	US-09-337-387-12	Sequence 12, Appli
43	2891.5	83.5	657	4	US-09-256-194-2	Sequence 2, Appli
44	2888.5	83.4	665	2	US-08-448-603A-32	Sequence 32, Appli
45	2888.5	83.4	665	3	US-09-134-075-32	Sequence 32, Appli

## ALIGNMENTS

RESULT 1  
US-08-448-603A-28  
Sequence 28, Application US/08448603A  
Patent No. 5864027  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
APPLICANT: Nakamura, Gerald R.  
TITLE OF INVENTION: HIV Envelope Polypeptides  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESSES:  
ADDRESS: McCutchen, Doyle, Brown & Enersen, LLP  
STREET: 3 Embarcadero Center  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,603A  
FILING DATE: 07-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/072,833  
FILING DATE: 07-JUN-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Halliday, Emily  
REGISTRATION NUMBER: 38903  
REFERENCE/DOCKET NUMBER: 14918-704  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-393-2000  
TELEFAX: 415-393-2286  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 850 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-448-603A-28  
Query Match 89.3%; Score 3091.5; DB 2; Length 850;  
Best Local Similarity 89.7%; Pred. No. 3e-250;  
Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;  
QY 2 EKLMTVYGVGPVWKEATTLFCASDAKAYDTEVHWVWATHACVPTDPNPOEVLLENTE 61  
|||||

```

Db      31 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNFOEIGLENVTE 90
Qy      62 HNNMKNMVEQWEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSBGMERG 120
Db      91 NFNMMKNMVEQWEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSBGMERG 150
Qy      121 EIKNCSFNITTSIRDEVEKAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXSIFEP 180
Db      151 EIKNCSFNITTSIRDKMKNKAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXSIFEP 210
Qy      181 IPIHYCAPAGFALLKCDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEVVI 240
Db      211 IPIHYCAPAGFALLKCDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEVVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPARAYTGEIIGDIRQAHCNL 300
Db      271 RSNFNSNATITIVOLNESVEINCTRPNNTRKSIHIGPARAYTGEIIGDIRQAHCNL 330
Qy      301 SRAKMDTLKQIYIKLREOPENKTIIVFNHSSGDDPEIYVMSFNGCEGFYCNSTOLFNST 360
Db      331 SSTRKMNNTLKQIYIKLREH-NKTIIVFNHSSGDDPEIYVMSFNGCEGFYCNSTOLFNST 389
Qy      361 MN-----NNTGSGNTEGNTITLPCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLL 415
Db      390 MNVYTNNTGSGNDT-GRNITLQCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLL 448
Qy      416 LTRDGINENGTETFRGGGDMRDNMRSEFYKYVVKIEPLGVAPTYKCKRRVYQREKRAV 475
Db      449 LTRDGG-NNSETEIFRGGGDMRDNMRSELYKYVVKIEPLGVAPTYKCKRRVYQREKRAV 507
Qy      476 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQQRMLOLTWGI 535
Db      508 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQQRMLOLTWGI 567
Qy      536 KQLOARVLAVERYLKQDQLLGIGCCSGKLICTAVPNWASWSNKSJDRINNMNTWEMER 595
Db      568 KQLOARVLAVERYLKQDQLLGIGCCSGKLICTAVPNWASWSNKSJDRINNMNTWEMER 627
Qy      596 EIDNYSIETITLIESONQOEKNEOELLEDKASLNMWPDITKMLWY 643
Db      628 EIDNYSIETITLIESONQOEKNEOELLEDKASLNMWPDITKMLWY 675

RESULT 2
US-09-134-075-28
; Sequence 28, Application US/09134075
; Patent No. 6042836
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Eneersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatsSEO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134.075
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,603
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily

```

```

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
;
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-134-075-28

Query Match: 89.3%; Score 3091.5; DB 3; Length 850;
Best Local Similarity 89.7%; Pred. No. 3e-250;
Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;

Qy      2 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNFOEIGLENVTE 61
Db      31 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNFOEIGLENVTE 90
Qy      62 HNNMKNMVEQWEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSBGMERG 120
Db      91 NFNMMKNMVEQWEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSBGMERG 150
Qy      121 EIKNCSFNITTSIRDEVEKAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXSIFEP 180
Db      151 EIKNCSFNITTSIRDKMKNKAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXSIFEP 210
Qy      181 IPIHYCAPAGFALLKCDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEVVI 240
Db      211 IPIHYCAPAGFALLKCDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEVVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPARAYTGEIIGDIRQAHCNL 300
Db      271 RSNFNSNATITIVOLNESVEINCTRPNNTRKSIHIGPARAYTGEIIGDIRQAHCNL 330
Qy      301 SRAKMDTLKQIYIKLREOPENKTIIVFNHSSGDDPEIYVMSFNGCEGFYCNSTOLFNST 360
Db      331 SSTRKMNNTLKQIYIKLREH-NKTIIVFNHSSGDDPEIYVMSFNGCEGFYCNSTOLFNST 389
Qy      361 MN-----NNTGSGNTEGNTITLPCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLL 415
Db      390 MNVYTNNTGSGNDT-GRNITLQCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLL 448
Qy      416 LTRDGINENGTETFRGGGDMRDNMRSEFYKYVVKIEPLGVAPTYKCKRRVYQREKRAV 475
Db      449 LTRDGG-NNSETEIFRGGGDMRDNMRSELYKYVVKIEPLGVAPTYKCKRRVYQREKRAV 507
Qy      476 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQQRMLOLTWGI 535
Db      508 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQQRMLOLTWGI 567
Qy      536 KQLOARVLAVERYLKQDQLLGIGCCSGKLICTAVPNWASWSNKSJDRINNMNTWEMER 595
Db      568 KQLOARVLAVERYLKQDQLLGIGCCSGKLICTAVPNWASWSNKSJDRINNMNTWEMER 627
Qy      596 EIDNYSIETITLIESONQOEKNEOELLEDKASLNMWPDITKMLWY 643
Db      628 EIDNYSIETITLIESONQOEKNEOELLEDKASLNMWPDITKMLWY 675

RESULT 3
US-09-492-739-28
; Sequence 28, Application US/09492739
; Patent No. 6331404
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
 STREET: 3 Embarcadero Center  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: PASEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/492,739  
 FILING DATE: 27-Jan-2000  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/134,075  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haliday, Emily  
 REGISTRATION NUMBER: 38903  
 REFERENCE/DOCKET NUMBER: 14918-704

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-393-2000  
 TELEFAX: 415-393-2286  
 TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 850 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
 US-09-492-739-28

Query Match 89.3%; Score 3091.5; DB 4; Length 850;  
 Best Local Similarity 89.7%; Pred. No. 3e-250;  
 Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;

QY 2 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61  
 DB 31 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90  
 QY 62 HNNMKNMVMQMDIISLMDOSLKPCKVLTPLCVTLNCKOV-WATNTNDSGEMRG 120  
 DB 91 HNNMKNMVMQMDIISLMDOSLKPCKVLTPLCVTLNCKOV-WATNTNDSGEMRG 150  
 QY 121 EIKKNSFITTSIRDEVOKEYALFYKLDVVPIDNNNTSVRLISCPTSVITQACPKISFEP 180  
 DB 151 EIKKNSFITTSIRKMKNEVALFYKLDVVPIDNNNTSVRLISCPTSVITQACPKISFEP 210  
 QY 181 IPIHCAGAPAILKCDKTFNGKPCXNVSTVQCTHGIRPVYSTOLLNGLAEDEVYI 240  
 DB 211 IPIHCAGAPAILKCDKTFNGKPCXNVSTVQCTHGIRPVYSTOLLNGLAEDEVYI 270  
 QY 241 RSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGRIIGIRAHONI 300  
 DB 271 RSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGRIIGIRAHONI 330  
 QY 301 SRAKNNDTLKOIVILREOFENKTIIVFNHSSGDEPEIVHSHFCGEFFYCNSLTOLFNSI 360  
 DB 331 SRAKNNDTLKOIVILREOFENKTIIVFNHSSGDEPEIVHSHFCGEFFYCNSLTOLFNSI 389  
 QY 361 WN-----NNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPIRGOIRCSSNITGL 415  
 DB 390 WN-----NNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPIRGOIRCSSNITGL 448  
 QY 416 LTRDGIENGTEIRPGGDMRDMRSSEFYKYVVKI EPLGAVATPKCRVRVQEKXAV 475  
 DB 449 LTRDGIENGTEIRPGGDMRDMRSSEFYKYVVKI EPLGAVATPKCRVRVQEKXAV 507  
 QY 476 GIGAVFLGAGSTMGASMTLTVQARLLLSGIIVQOQNNLLRAIEAQORMLQLTWGI 535

DB 508 GIGAVFLGAGSTMGASMTLTVQARLLLSGIIVQOQNNLLRAIEAQORMLQLTWGI 567  
 QY 536 KOLARVLAVERYLGDQOLGIMGCSGKICTAVPMWASNSKSLDRIMNMTMWER 595  
 DB 568 KOLARVLAVERYLGDQOLGIMGCSGKICTAVPMWASNSKSLDRIMNMTMWER 627  
 QY 596 EIDNVTSEIYTLIEESQNOQEKNEOELELDKMASLWNPFDITWLMY 643  
 DB 628 EIDNVTSEIYTLIEESQNOQEKNEOELELDKMASLWNPFDITWLMY 675

RESULT 4  
 US-08-022-835-6  
 Sequence 6, Application US/08022835  
 Patent No. 5420030

GENERAL INFORMATION:  
 APPLICANT: Reitz Jr., Marvin S.  
 APPLICANT: Franchini, Genoveffa  
 APPLICANT: Markham, Phillip D.  
 APPLICANT: Gallo, Robert C.  
 APPLICANT: Lori, Franco C.  
 APPLICANT: Popovic, Mikulas  
 APPLICANT: Garner, Suzanne  
 TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES  
 THEREOF  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN, DABY & CUSHMAN  
 STREET: Eleventh floor, 1615 L. Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20036-5601

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/022,835  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/599,491  
 FILING DATE: 17-OCT-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Scott, Watson T.  
 REGISTRATION NUMBER: 26,581  
 REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 861-3000  
 TELEFAX: (202) 822-0944  
 TELEX: 6714627 CUSH  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 855 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-022-835-6

Query Match 88.7%; Score 3069.5; DB 1; Length 855;  
 Best Local Similarity 88.1%; Pred. No. 2.1e-248;  
 Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

QY 2 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61  
 DB 31 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90  
 QY 62 HNNMKNMVMQMDIISLMDOSLKPCKVLTPLCVTLNCKOV-----WATNTNDSG 115  
 DB 91 HNNMKNMVMQMDIISLMDOSLKPCKVLTPLCVTLNCKOV-----WATNTNDSG 150

Qy	116	IMEREIEIKCSNITITSIDEVOKEALYKLDVIBDNN-NSYSILISCDTIVIQACP	174
Db	151	MYGSEBMKCSNITITNIRGKVOKEALYKLDIAIDNNSNRKYLISNTSVITIQACP	210
Qy	175	KISEPPIPHHYCAPAGFAILKCNKDTFNGKPGCKSNSTVQCTHYGIRPVVSTOLLNGSLA	234
Db	211	KVSFEPPIPHYCAPAGFAILKCKDKKFNKPGCTNVSIVQCHGIRPVVSTOLLNGSLA	270
Qy	235	EEBVYIRSDNFNNAKTIIVOLKESEYAIACTRPENNTRKSIHIGPRAFYTTGEIIGDIR	294
Db	271	EEBVYIRSANFADNMKVIIIVOLNESYEINCTRPNNTTRKSIHIGPRAFYTTGEIIGDIR	330
Qy	295	QAHCHISRAKNMDTLKQIYIKIREQENKTIYVNHSSGDPPEIVMHSFNCEGFEFFYCNST	354
Db	331	QAHCHLSRAKNMDTLNKIYIKIREQGNKTIYFKHSSGDPPEIVTHSFNCGGFEFFYCNST	390
Qy	355	QLFNSTNNNTGESSNTEGNITTLPCRIOIIMMOGEVKAMAPPIRGDIRCSSNITGL	414
Db	391	QLFNSTNNVTESSNNTVENNTTLPCRIOIIMMOGEVRAMAPPIRGDIRCSSNITGL	450
Qy	415	LITRDGGINENGETEIPRPGGDMRDMRSEFYKYKVKIIEPLGVAFTPKCKRRVVOREKRA	474
Db	451	LITRPGGPEDNKTEVEFRPGGDMRDMRSELYKVKVKEIETPLGVAFTPKARRVVOREKRA	510
Qy	475	VGIGAVPFGFLGAGSGTMAASMTLTVQARLLSGIVOOQNNLIRAIIEAQORMLQITVWG	534
Db	511	VGIGAVPFGFLGAGSGTMAAAMTLTVQARLLSGIVOOQNNLIRAIIEAQOHLQITVWG	570
Qy	535	IKOLQARVLAVERYLADQOLLGIGWCSGKRLICTTAPVPMNASWMSKSIDRIANNMTWME	594
Db	571	IKOLQARVLAVERYLADQOLLGIGWCSGKRLICTTAPVPMNASWMSKSLINKIMDNWTIEMD	630
Qy	595	REIDNYSITVLIESONQOEENBELEELDQWASLMMWEDITNNLMY	643
Db	631	REINNTYSITVLIESONQOEENBELEELDQWASLMMWEDITNNLMY	679

RESULT 5  
 US-08-388-809-6  
 Sequence 6, Application US/08388809  
 Patent No. 557600  
 GENERAL INFORMATION:  
 APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
 APPLICANT: GENOVEEPA, MARKHAM, PHILIP D. GALLO, ROBERT  
 APPLICANT: C.; LOETI, FRANCO C.; POPOVIC, MIKULAS; AND  
 APPLICANT: GARTNER, SUZANNE  
 TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
 TITLE OF INVENTION: AND USES THEREOF  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORGAN & FINNEGAN  
 STREET: 345 PARK AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10154  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY DISK, 3.5"  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WORDPERFECT 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/388,809  
 FILING DATE: 15-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/022,835  
 FILING DATE: 25-FEB-1993  
 APPLICATION NUMBER: US 07/599,491  
 FILING DATE: 17-OCT-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LESLIE A. SERINIAN  
 REGISTRATION NUMBER: 35,353  
 REFERENCE/DOCKET NUMBER: 2026-4092052

```

1 TELECOMMUNICATION INFORMATION
2 TELEPHONE: (212) 758-4800
3 TELEFAX: (212) 751-6849
4 TELEX: 421792
5
6 INFORMATION FOR SEQ ID NO: 6:
7 SEQUENCE CHARACTERISTICS:
8 LENGTH: 855 amino acids
9 TYPE: amino acids
10 TOPOLOGY: linear
11
12 MOLECULE TYPE: protein
13
14 US-08-388-809-6

```

Query Match	88.7%;	Score 3069.5;	DB 1;	Length 855;
Best Local Similarity	88.1%;	Pred. No. 2.1e-248;		
Matches 572;	Conservative 27;	Mismatches 43;	Indels 7;	Gaps 2.

```

QY 2 KKLWTVYVYGVPMVEATTTTLEFCADADAYPDEVHNVATHACVPTDPMPOEVLNENVE 61
Db 31 EKLWTVYVYGVPMVEATTTTLEFCADBRKAYDDEVHNVATHACVPTDPMPOEVLKNVE 90
QY 62 HFNMKNNMVEQMOEDIISLMDQSLKPCVKLTPLCCVTLNCKDV-----NATNTNDSRG 115
Db 91 NFNMKNNMVMQMHEDIISLMDQSLKPCVKLTPLCCVTLNCTDLRATNAGNDNTNTSSSRG 150
QY 116 TMERREINCSFNITTSIRDEQKEXYALFYKLDVYPIDNN-NTSYRLISCDTSVLTQACP 174
Db 151 MWGGSEEMKNCSENTITNIRGKQKEXALFYKLDIAPIDNNSNNRRLISCSNTSVLTQACP 210
QY 175 KISFEPRIHYCAPAFALILKNDTFNGKGCCKNVSTVQCTHGIRPVVSTOLLNGLSJA 234
Db 211 KVSFEPRIHYCAPAFALILKCKDKKFKNGKGTJVSTVQCTHGIRPVVSTOLLNGLSJA 270
QY 235 EBEVIRSDNFNTNNAKTIIIVOLKESVEINCTRPNNNTRKS IHIGGRAYTTTGEIIGDIR 294
Db 271 EBEVIRSRANPDANKVIIIVOLNESEVEINCTRPNNNTRKS IHIGGRAYTTTGEIIGDIR 330
QY 295 OAHCHISAQKNDITLKOIYIKLREQPEKNTIYFNHSSGSDPEIYNHSPNCEGFEFYCNST 354
Db 331 OAHCHISAQKNDITLNTKIYIKLREQPFKNITVFKHSSGSDPEIYHSPNCEGFEFYCNST 390
QY 355 QLFNSTWNNNTSGSNTEGNTITTLPCRIKQIINNMQEVGKAMYAPPIRQOIRCSSNITGL 414
Db 391 QLFNSTWNNNTSGSNTEGNTITTLPCRIKQIINNMQEVGRAMYAPPIRQOIRCSSNITGL 450
QY 415 LITRGGINGNGETLEFRPGGDMRNMWSEPKYKVVYKIEPGLVAPTKCKRRPVQOEKXA 474
Db 451 LITRGGEGEDNTEVFRPGGDMRNMWSELPKYKVKIEPGLVAPTKCKRRPVQOEKXA 510
QY 475 VGIQAVPFLGPGAAGSTWGAASMTLTVQARLLISGIVQOQNNLLRAIEAQOERMLQJTWG 534
Db 511 VGIQAVPFLGPGAAGSTWGAASMTLTVQARLLISGIVQOQNNLLRAIEAQOHLQJTWG 570
QY 535 IKOLQARVLAVERYIGDOOLIGWCSGKLICTTAVPMNMASSNKSJLBIINNNMTMME 594
Db 571 IKOLQARVLAVERYIRDOOLLGIWCSGKLICTTAVPMNMASSNKSJLKNIMDMTWIEMD 630
QY 595 REIDNYTSEIYTLIESONQOEKNOELLELDKMASLNMWPIITWMLY 643
Db 631 REINNYTSLIYLSIESONQOEKNOELLELDKMASLNMWPIITKMLY 679

```

RESULT 6  
 US-08-6647-714-6  
 ; Sequence 6, Application US/08647714  
 ; Patent No. 5869313  
 ;  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
 ; APPLICANT: GENOVESEFA; MARKHAM, PHILIP D. GALLO, ROBERT  
 ; APPLICANT: C.; LORI, FRANCO C.; POBOVIC, MIKLAS; AND  
 ; APPLICANT: GARNER, SUZANNE  
 ;  
 ; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
 ;  
 ; TITLE OF INVENTION: AND USES THEREOF  
 ;  
 ; NUMBER OF SEQUENCES: 6



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK, 3.5"
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,714
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/388,809
FILING DATE: 15-FEB-1995
APPLICATION NUMBER: US 08/022,835
FILING DATE: 25-FEB-1993
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: LESLIE A. SERUNIAN
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4092US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-647-714-6

Query Match      88.7%; Score 3069.5; DB 2; Length 855;
Best Local Similarity 88.1%; Pred. No. 2,1e-248;
Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

QY 2 EKLWTVVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWTVVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 115
DB 91 NFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCTDLRNATNGNDTNTSSRG 150
QY 116 TMRGEIKNCSFNTTTSIRDEVQKRYALFYKLDVVPIDNN-NTSYRLISCTSVTQACP 174
DB 151 WVGGEKNCSPNTTNRGRVQKRYALFYKLDIAPIDNNSNRRLISCTSVTQACP 210
QY 175 KISFEPIPIHCAFGAIFALKCNKDTFNGKPCCKVSTVOCTHGRVPVSTOLLNGSLA 234
DB 211 KVSFEPIPIHCAFGAIFALKCKKCKKNGKPCCTVSTVOCTHGRVPVSTOLLNGSLA 270
QY 235 EEEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTKSIHIGRAFYTTGELIGDIR 294
DB 271 EEEVVISRDNFTNNAKTIIVQLNESVEINCTRPNNNTKSIHIGRAFYTTGELIGDIR 330
QY 295 QAHNINISAKKNDPLTKOIVILKREPFENKTIIVFNHSSGDEBEIVHSHNGEPEFYCNST 354
DB 331 QAHNINISAKKNDPLTKIVILKREPFENKTIIVFKISSGDEBEIVHSHNGEPEFYCNST 390
QY 355 QLFNSTWNNTEGSGNTTGTITLLPCRKOIINMOEYKAMVAPPPIGQIRCSNITGL 414
DB 391 QLFNSTWNNTEGSGNTTGTITLLPCRKOIINMOEYKAMVAPPPIGQIRCSNITGL 450
QY 415 LTRDGGINGENGTETIFRPGGGDMRDNWSSEFYKRVVXIIEPLGVAPTKCKRNVVQREKRA 474
DB 451 LTRDGGINGENGTETIFRPGGGDMRDNWSSEFYKRVVXIIEPLGVAPTKCKRNVVQREKRA 510

```

```

475 VGIGAVFLGFIAGSGTWGASMTLTVQARLLLSGIVOOQNNLRAIEAQOQMLQLTWVG 534
511 VGIGAVFLGFIAGSGTWGASMTLTVQARLLLSGIVOOQNNLRAIEAQOQMLQLTWVG 570
QY 535 IKOLARVLAVERYLDQOLLGIGWCSGKLICTAVPNNASWSNKLDRINNNNTWMEWE 594
DB 571 IKOLARVLAVERYLDQOLLGIGWCSGKLICTAVPNNASWSNKLDRINNNNTWMEWE 630
QY 595 REINDYTSIYTLIBESQNOQEKNOEELIEDKASLWNPFDITNWLWY 643
DB 631 REINDYTSIYTLIBESQNOQEKNOEELIEDKASLWNPFDITNWLWY 679

RESULT 7
US-07-956-483-14
Sequence 14, Application US/07956483
Patent No. 6261799
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF INVENTIONS: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 2213-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,483
FILING DATE: 31-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/19742
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 05392
FILING DATE: 02-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 017753-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match      87.4%; Score 3027.5; DB 3; Length 855;
Best Local Similarity 87.2%; Pred. No. 6.9e-245;
Matches 571; Conservative 33; Mismatches 34; Indels 17; Gaps 7;

QY 2 EKLWTVVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWTVVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 115
DB 91 NFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCTDLRNATNGNDTNTSSRG 150

```

Qy	116	TMERGEJHNSFNITTSIRDEVOKEALYKLDVPIEDNNNSYALSCDSVTOACPK	175
Db	151	KMEGEGHTNCSFPIITTSIRSQVKEALYKLDVPIED--NTSYLLMCSNVSYTOACPK	208
Qy	176	ISPEPIPIHCABAGFALLKCDKDTENGKPCKNVSYVQCHGIRPVVSTOLLNGSLAE	235
Db	209	VSEPEPIPIHCA--RWFPAIINCNNKKENGSGPCTNVSTVQCHGIRPVVSTHLLNGSLAE	267
Qy	236	EEVYIREBNTNNAKTIIVOLKESVINECTRPNNTNRSIHIGPERAYTTEIGIRQ	299
Db	268	EEVYIREBNTNNAKTIIVOLKESVINECTRPNNTNRSIHIGPERAYTTEIGIRQ	322
Qy	296	AHCNISPAKANNITLKQIVIKLREOFENKTIIVENHSGGPEIYVMSFNCGEFPCYCNSTQ	355
Db	328	AHCNISPAKANNITLKQIVIKLREOFENKTIIFNRSGGPEIYVMSFNCGEFPCYCNSTQ	387
Qy	356	LEFSTMMNNNEGSNNTEGN--TTLPRIOIIMMOEUCKAMYPAPIRGOIRGCSNITGL	415
Db	388	LEFSTMM--NGTEGSNNNGDNTLLPRICEIIMMOEUCKAMYPAPIRGOIRGCSNITGL	446
Db	415	LLTRDGINENG-----TEIFRPGGGDMRDNRSFEYKXKVKLEPLGVAATPKKRRVYQ	465
Db	447	LLTRDGSNSKGSKNENETEIFRPGGDMRDNRSLEYKXKVKLEPLGVAATPKKRRVYQ	506
Qy	470	REKRAVY--IAVPLGFGAAGSTMGASMTLLVQARLLSGIVOOONMLLAIRAQOML	528
Db	507	REKRAVSTIAPMLGFGAAGSTMGATSMTLVQARLLSGIVOOONMLLAIRAQOHL	566
Qy	529	QLTWGIKQOLQARVILVERLYGDOQLLGIMGCSGKLICTAYPMNASMSNKSLDRIWNMM	588
Db	567	QLTWGIKQOLQARVILVERLYRDOQLLGIMGCSGKLICTTAYPMNTSMNSKSLDKINGMM	622
Qy	589	TYMWEEMELINYSLEYTLLEESQNOEQNEDELLBDMKASLMMWEPITWMLAY	643
Db	627	TYMWEEMELINYSLEYTLLEESQNOEQNEDELLBDMKASLMMWEPITWMLAY	681

```

? INFORMATION FOR SEQ ID NO: 30:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 857 amino acids
?
? TYPE: amino acid
?
? STRANDEDNESS: single
?
? TOPOLOGY: linear
?
? US-08-448-603A-30

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134.075
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/448,603
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-134-075-30

```

```

Query Match      86.6%; Score 2999; DB 3; Length 857;
Best Local Similarity 85.4%; Pred. No. 1.7e-242;
Matches 556; Conservative 41; Mismatches 44; Indels 10; Gaps 3;

```

```

QY 3 KLMVTVVYGVVPMWEATTTTLFCASDAKAYDTEVHVNVMTACVPTDPNPOEVLENVTEH 62
DB 32 KLMVTVVYGVVPMWEATTTTLFCASDAKAYDTEVHVNVMTACVPTDPNPOEVLENVTEH 91
QY 63 FNMVKNVNEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 121
DB 92 FNMVKNVNEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 151
QY 122 IKNSFNITTSIRDEVQKVALFYKLDVVPIDN-----NNTSYRLISCOTSVITQACP 174
DB 152 IKNSFNITTSIRDEVQKVALFYKLDVVPIDN-----NNTSYRLISCOTSVITQACP 211
QY 175 KISFPIPIHYCAPAGFALILKCNKDTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFPIPIHYCAPAGFALILKCNKDTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 294
DB 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331
QY 295 QAHCHISRAKNDTLKQIYIKLREOPENKTIIVFNHSSGGDPEI VMSFNCGEFFYCNST 354
DB 332 QAHCHISRAKNDTLKQIYIKLREOPENKTIIVFNHSSGGDPEI VMSFNCGEFFYCNST 391
QY 355 QLFNSTNNNN--TEGSSNTEGNTITLPCRIKQIIMMOBVGKAMAPPRGQIRROSSNT 412
DB 392 QLFNSTNNNN--TEGSSNTEGNTITLPCRIKQIIMMOBVGKAMAPPRGQIRROSSNT 451
QY 413 GLLLTRDGINENGTEIFRPGGDMRDWRSEFYKYKVKVIEPLGVAFTKCRARVOREK 472
DB 452 GLLLTRDGINENGTEIFRPGGDMRDWRSEFYKYKVKVIEPLGVAFTKCRARVOREK 511
QY 473 RAVGIGAVFLGLGAAGSTMGAASMTLTVQARLLISGIYOQNNILRAIEAOQRMQLDTV 532
DB 512 RAVGIGAVFLGLGAAGSTMGAASMTLTVQARLLISGIYOQNNILRAIEAOQRMQLDTV 571
QY 533 WGIKLOLQAVLAVERYLDQOOLIGIWGSGKLIICCAVAPVMAWSKSLDRIMNMNTWME 592
DB 572 WGIKLOLQAVLAVERYLDQOOLIGIWGSGKLIICCAVAPVMAWSKSLDRIMNMNTWME 631
QY 593 WEREIDNTSEIYTLIEESQNOQEKNEOBLLBLDKWASLMMWFDITNMLMY 643

```

```

DB 632 WEREIDNTSEIYTLIEESQNOQEKNEOBLLBLDKWASLMMWFDITNMLMY 682

```

```

RESULT 10
US-09-492-739-30
Sequence 30, Application US/09492739
Patent No. 6331404
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
Nakanura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: McCutchen, Doyle, Brown & Eversen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-492-739-30

```

```

Query Match      86.6%; Score 2999; DB 4; Length 857;
Best Local Similarity 85.4%; Pred. No. 1.7e-242;
Matches 556; Conservative 41; Mismatches 44; Indels 10; Gaps 3;

```

```

QY 3 KLMVTVVYGVVPMWEATTTTLFCASDAKAYDTEVHVNVMTACVPTDPNPOEVLENVTEH 62
DB 32 KLMVTVVYGVVPMWEATTTTLFCASDAKAYDTEVHVNVMTACVPTDPNPOEVLENVTEH 91
QY 63 FNMVKNVNEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 121
DB 92 FNMVKNVNEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 151
QY 122 IKNSFNITTSIRDEVQKVALFYKLDVVPIDN-----NNTSYRLISCOTSVITQACP 174
DB 152 IKNSFNITTSIRDEVQKVALFYKLDVVPIDN-----NNTSYRLISCOTSVITQACP 211
QY 175 KISFPIPIHYCAPAGFALILKCNKDTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFPIPIHYCAPAGFALILKCNKDTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 294
DB 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331

```

QY 295 QAHNCISRAKNDTLKQIVILKREOFENKTIIVFNHSSGDEPEIVMHSNCEGFEFFYCNST 354  
 Db 332 QAHNCISRTDWNNTLQIVKREOFENKTIIVFNHSSGDEPEIVMHSNCEGFEFFYCNST 391  
 QY 355 QLFNSTMN--TEGSNTEGNTITLPCRIOIIMMOEYKAMAPPIRQIRCSNIT 412  
 Db 392 QLFNSTMN--TEGSNTEGNTITLPCRIOIIMMOEYKAMAPPIRQIRCSNIT 451  
 QY 413 GLILTRDGINENGTEIFRPGGDMRDWRSEFFYKVKXIEPLGVAFTKCRKRVOREK 472  
 Db 452 GLILTRDGINENGTEIFRPGGDMRDWRSEFFYKVKXIEPLGVAFTKCRKRVOREK 511  
 QY 473 RAVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQOQMLQITV 532  
 Db 512 RAVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQOQMLQITV 571  
 QY 533 WGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMAWSNKSLEDIRNNMTWME 592  
 Db 572 WGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMAWSNKSLEDIRNNMTWME 631  
 QY 593 WEREIDNTSEIYTLIEESONQOEKNEBELLEDKMAISLMMWPDITNMLWY 643  
 Db 632 WEREIDNTSEIYTLIEESONQOEKNEBELLEDKMAISLMMWPDITNMLWY 682

RESULT 11  
 US-08-127-499A-14  
 : Sequence 14, Application US/08127499A  
 : Patent No. 5510264  
 : GENERAL INFORMATION:  
 : APPLICANT: VAN ALSTYNE, Diane  
 : TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
 : TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
 : NUMBER OF SEQUENCES: 40  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Foley & Lardner  
 : STREET: 3000 K Street, N.W., Suite 500  
 : CITY: Washington  
 : STATE: D.C.  
 : COUNTRY: USA  
 : ZIP: 20007-5109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/127,499A  
 : FILING DATE: 28-SEP-1993  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: BENT, Stephen A.  
 : REGISTRATION NUMBER: 29,768  
 : REFERENCE/DOCKET NUMBER: 51916/102/INBI  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (202) 672-5300  
 : TELEFAX: (202) 672-5399  
 : TELEX: 904136  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 861 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS:  
 : TOPOLOGY: unknown  
 : US-08-127-499A-14

Query Match 86.4%; Score 2992.5; DB 1; Length 861;  
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;  
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;  
 QY 2 EKLMTVTVYGVVWKEATITLFCASDAKAYTEVHNWMTACVPTDPNPOEVLNVATE 61  
 Db 32 EKLMTVTVYGVVWKEATITLFCASDAKAYTEVHNWMTACVPTDPNPOEVLNVATE 91

QY 62 HFNMMKNNMTEQMOEDIIISLMDQSLKPCVLTPLCVTLNCKDY--NATNT-----TNSEG- 115  
 Db 92 NFNMMKNNMTEQMOEDIIISLMDQSLKPCVLTPLCVTLNCKDY--NATNT-----TNSEG- 151  
 QY 116 -TMRGEIKNCSFNITTSIDEVQKEYALFYKLDVVIDNNNTSYRLISCDTSVITQACP 174  
 Db 152 MMEKEGEEKNCSFNISISIRGKVQKEYAFPKLDIIPIDIDTTSYRLISCDTSVITQACP 211  
 QY 175 KISPEPIHYCA PAGAILKNDKTFNGKPCKNSTVQCHGIRPVYSTOLLNLSLA 234  
 Db 212 KISPEPIHYCA PAGAILKNDKTFNGKPCKNSTVQCHGIRPVYSTOLLNLSLA 271  
 QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNNTSKSIH--GPGAFYTTGELIGD 292  
 Db 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNNTSKSIH--GPGAFYTTGELIGD 330  
 QY 293 IROAHNCISRAKNDTLKQIVILKREOF--ENKTIIVFNHSSGDEPEIVMHSNCEGFEFFYC 351  
 Db 331 IROAHNCISRAKNDTLKQIVILKREOF--ENKTIIVFNHSSGDEPEIVMHSNCEGFEFFYC 390  
 QY 352 NSTOLFENSTMN--NTEGSNTEG--NTITLPCRIOIIMMOEYKAMAPPIRQIRCSNIT 407  
 Db 391 NSTOLFENSTMN--NTEGSNTEG--NTITLPCRIOIIMMOEYKAMAPPIRQIRCSNIT 450  
 QY 408 SSNITGLILTRDGINENGTEIFRPGGDMRDWRSEFFYKVKXIEPLGVAFTKCRKRV 467  
 Db 451 SSNITGLILTRDGINENGTEIFRPGGDMRDWRSEFFYKVKXIEPLGVAFTKCRKRV 510  
 QY 468 VOREKRAVIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQOQMLQITV 527  
 Db 511 VOREKRAVIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQOQMLQITV 570  
 QY 528 IOLTWGIRKOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMAWSNKSLEDIRNNMTWME 587  
 Db 571 IOLTWGIRKOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMAWSNKSLEDIRNNMTWME 630  
 QY 588 MTEMEREIDNTSEIYTLIEESONQOEKNEBELLEDKMAISLMMWPDITNMLWY 643  
 Db 631 MTEMEREIDNTSEIYTLIEESONQOEKNEBELLEDKMAISLMMWPDITNMLWY 686

RESULT 12  
 US-08-482-847-14  
 : Sequence 14, Application US/08482847  
 : Patent No. 5556757  
 : GENERAL INFORMATION:  
 : APPLICANT: VAN ALSTYNE, Diane  
 : TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPE SITES FOR  
 : TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
 : TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
 : NUMBER OF SEQUENCES: 40  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Foley & Lardner  
 : STREET: 3000 K Street, N.W., Suite 500  
 : CITY: Washington  
 : STATE: D.C.  
 : COUNTRY: USA  
 : ZIP: 20007-5109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/482,847  
 : FILING DATE: 07-JUN-1995  
 : CLASSIFICATION: 514  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 06/127,499  
 : FILING DATE: 28-SEP-1993  
 : ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 51916/104/INBI  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 672-5300  
 TELEFAX: (202) 672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 US-08-482-847-14

Query Match 86.4%; Score 2992.5; DB 1; Length 861;  
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;  
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTY 61
32 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTY 91
62 HNNMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVTLNCKDV--NATNT---TNDSEG- 115
92 NFNMMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATVNTNSNTSSSGE 151
116 -TMRGEIKNCSFNITTSIRBEVQKVALFYKLDVVPIDNNNTSYRLISCTSVITQACP 174
152 MMEKEGELKNCSFNITTSIRBEVQKVALFYKLDIIPIDNDTSTYRLISCTSVITQACP 211
175 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCKAVSTVOCTHGIRPVVSTOLLNGSLA 234
212 KVSFEPIPIHYCAPAGFAILKCNKDTFNGKPCKAVSTVOCTHGIRPVVSTOLLNGSLA 271
235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPRAFYTTGEIIGD 292
272 EEEVIRSANFTDNAKTIIVOLQNSVEINCTRPNNNTRKSIIRIORPRAFTVIGK-IGN 330
293 IROAHCHNSRAKMNLTLOIYIKLREOF-EKKTIVFNHSSGGDEIIVHNSNGCEFPYC 351
331 MROAHCHNSRAKMNLTLOIYIKLREOF-EKKTIVFNHSSGGDEIIVHNSNGCEFPYC 390
352 NSTOLFNSTWNN--NTGSGNTEG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRC 407
391 NSTOLFNSTWNN--NTGSGNTEG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRC 450
408 SSNITGLLLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVYKIEPLGVAFTKCKRRV 467
451 SSNITGLLLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVYKIEPLGVAFTKCKRRV 510
468 VOREBRAVIGAVPLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQQHL 527
511 VOREBRAVIGAVPLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQQHL 570
528 LQLTWGKIQOLQARLAVERYLGDQQLGIMGCSKLLCTTAVPNNASMSKSLDRIRVNN 587
571 LQLTWGKIQOLQARLAVERYLGDQQLGIMGCSKLLCTTAVPNNASMSKSLDRIRVNN 630
588 MTWMEWEIEDNYTSEIYTLIEESQOQEKNOEQLLELDKVASLWMMEDITWMLVY 643
631 MTWMEWEIEDNYTSEIYTLIEESQOQEKNOEQLLELDKVASLWMMEDITWMLVY 686

```

RESULT 13  
 US-07-956-483-10  
 Sequence 10, Application US/07956483  
 Patent No. 6261799  
 GENERAL INFORMATION:  
 APPLICANT: KIENY, Marie-Paule  
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
 TITLE OF INVENTION: gp160 VARIANT  
 NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P O Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/956,483  
 FILING DATE: 31-DEC-1992  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO 92/19742  
 FILING DATE: 12-NOV-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 91 05392  
 FILING DATE: 02-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 01753-005  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-07-956-483-10

Query Match 86.4%; Score 2992.5; DB 3; Length 861;  
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;  
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTY 61
32 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTY 91
62 HNNMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVTLNCKDV--NATNT---TNDSEG- 115
92 NFNMMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATVNTNSNTSSSGE 151
116 -TMRGEIKNCSFNITTSIRBEVQKVALFYKLDVVPIDNNNTSYRLISCTSVITQACP 174
152 MMEKEGELKNCSFNITTSIRBEVQKVALFYKLDIIPIDNDTSTYRLISCTSVITQACP 211
175 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCKAVSTVOCTHGIRPVVSTOLLNGSLA 234
212 KVSFEPIPIHYCAPAGFAILKCNKDTFNGKPCKAVSTVOCTHGIRPVVSTOLLNGSLA 271
235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPRAFYTTGEIIGD 292
272 EEEVIRSANFTDNAKTIIVOLQNSVEINCTRPNNNTRKSIIRIORPRAFTVIGK-IGN 330
293 IROAHCHNSRAKMNLTLOIYIKLREOF-EKKTIVFNHSSGGDEIIVHNSNGCEFPYC 351
331 MROAHCHNSRAKMNLTLOIYIKLREOF-EKKTIVFNHSSGGDEIIVHNSNGCEFPYC 390
352 NSTOLFNSTWNN--NTGSGNTEG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRC 407
391 NSTOLFNSTWNN--NTGSGNTEG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRC 450
408 SSNITGLLLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVYKIEPLGVAFTKCKRRV 467
451 SSNITGLLLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVYKIEPLGVAFTKCKRRV 510

```

Qy 468 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVOOONLLRAIEAOQM 527  
Db 511 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVOOONLLRAIEAOQH 570  
Qy 528 LQLTWVGIRKOLQARVLAVERYLGDQOLLGIVGSGSKLICCTAVPWNASWSNKSIDRIWNN 587  
Db 571 LQLTWVGIRKOLQARVLAVERYLKDQOLLGIVGSGSKLICCTAVPWNASWSNKSLEQIWN 630  
Qy 588 MTWMEWREIDNTSEIYTLIEESONOEKNEBELLEDKASLMMNFDTNMLMY 643  
Db 631 MTWMEWREIDNTSEIYTLIEESONOEKNEBELLEDKASLMMNFDTNMLMY 686

RESULT 14  
US-08-472-240A-1  
Sequence 1, Application US/08472240A  
Patent No. 6284248  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472.240A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,483  
FILING DATE: 31-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 017753-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-6620  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..854  
US-08-472-240A-1

Query March 86.4%; Score 2992.5; DB 3; Length 861;  
Best Local Similarity 86.7%; Pred. No. 5.9e-242;  
Matches 569; Conservative .32; Mismatches 40; Indels 15; Gaps 8;  
Qy 2 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 61  
Db 32 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLVNTE 91  
Qy 62 HFNWKNMNVQOMQEDTISLWDSIKPCVKLTPLCTVLANQDV-NATNT---TNDSEG- 115  
Db 92 NFNWKNMNVQOMQEDTISLWDSIKPCVKLTPLCTVLANQDV-NATNT---TNDSEG- 151

Qy 116 -TWGEIKNCSFNITTSIRDEYOKEXALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 174  
Db 152 MMKEGEBIKNCSFNITTSIRGKXVQKEAFYKLDIIPIDDTTSYRLISCDTSVITQACP 211  
Qy 175 KISEEPIPIHYCAPAGFAIILKCDKTPNGKPCANSTVOCTGIRPVSTOLLNSLA 234  
Db 212 KVSFEPIPIHYCAPAGFAIILKCNKTFNGTGPCTNVSTVOCTGIRPVSTOLLNSLA 271  
Qy 235 EEEVVRSDNFTNAKTIIVQLKESVEINCTRPNNNTKSIH--GGRGFYTTGEIIGD 292  
Db 272 EEEVVRSDNFTNAKTIIVQLKESVEINCTRPNNNTKSIH--GGRGFYTTGEIIGD 330  
Qy 293 IRQAHCNISRAKNNDTLKQIVILKEQF--ENKTIIVNHSAGDEIYVMSFNGEPEFYC 351  
Db 331 MRQAHCNISRAKNNDTLKQIVILKEQF--ENKTIIVNHSAGDEIYVMSFNGEPEFYC 390  
Qy 352 NSTOLFNSITWNN--NTGSSNNTBG-NTTLPCRIKQIIMMOBVGKAMTAPPIRGQIRC 407  
Db 391 NSTOLFNSITWNN--NTGSSNNTBG-NTTLPCRIKQIIMMOBVGKAMTAPPIRGQIRC 450  
Qy 408 SSNTGLLTRDGINENGTEIFRPGGDMRDMRSFYYKYVKTIEPLGVAPTKGRV 467  
Db 451 SSNTGLLTRDGINENGTEIFRPGGDMRDMRSFYYKYVKTIEPLGVAPTKGRV 510  
Qy 468 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVOOONLLRAIEAOQM 527  
Db 511 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVOOONLLRAIEAOQH 570  
Qy 528 LQLTWVGIRKOLQARVLAVERYLGDQOLLGIVGSGSKLICCTAVPWNASWSNKSIDRIWNN 587  
Db 571 LQLTWVGIRKOLQARVLAVERYLKDQOLLGIVGSGSKLICCTAVPWNASWSNKSLEQIWN 630  
Qy 588 MTWMEWREIDNTSEIYTLIEESONOEKNEBELLEDKASLMMNFDTNMLMY 643  
Db 631 MTWMEWREIDNTSEIYTLIEESONOEKNEBELLEDKASLMMNFDTNMLMY 686

RESULT 15  
US-08-472-240A-7  
Sequence 7, Application US/08472240A  
Patent No. 6284248  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472.240A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,483  
FILING DATE: 31-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 017753-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-6620  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..831  
 US-08-472-240A-7

Query Match 86.4%; Score 2992.5; DB 3; Length 861;  
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;  
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

QY 2 EKLMTVTVYGVVPMWEATTTTLFCASDAKADYDEVNWNATHACVPTDPNPOEVLENTE 61
DB 32 EKLMTVTVYGVVPMWEATTTTLFCASDAKADYDEVNWNATHACVPTDPNPOEVLENTE 91
62 HFNMKKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
92 NFNMKKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSSE 151
116 -TMRGEIKNCSFNITTSIRDEVQKEVALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 174
152 MMEKEGIKNCSFNISTIRGVQKEVAFYKLDIIPIDNDTTSYTLTSCNTSVITQACP 211
175 KISFPIPIHICAPAGFAILKNDKTFNGKGPCKAVSTVQCTHGIRPVYSTOLLNGSLA 234
212 KVSFPIPIHICAPAGFAILKCNKTFNGTGVCTVSTVQCTHGIRPVYSTOLLNGSLA 271
235 EEEVIVRSDFNFTNNAKTIIVOLKESVVEINCTRPNNNTRKSIH--GPGRAFTTGEITGD 292
272 EEEVIVIRSANFTDNAKTIIVOLNGSVEINCTRPNNNTRKSIH--GPGRAFTTGEITGD 330
293 IROAHCNISRAKMDTLKQIVIKLEQF-ENKTIIVNHSGGDPEIIMHSFNCGEFPFYC 351
331 MROAHCNISRAKMDTLKQIVIKLEQF--GNNKTIIFKQSSGGDPEIIVHSGGEFPFYC 390
352 NSTOLFNSTWNTN--NTGSSNTEG--NTITLPCRIKQIIMWQEVGKAMYAPPIRGQIRG 407
391 NSTOLFNSTWNTN--NTGSSNTEG--NTITLPCRIKQIIMWQEVGKAMYAPPIRGQIRG 450
408 SSNITGLLITRDGGINENGTEIFRPGGDMRDNRSEFYKVKVLEPIGVAPTCKKRV 467
451 SSNITGLLITRDGGINENGTEIFRPGGDMRDNRSEFYKVKVLEPIGVAPTCKKRV 510
468 VOREKRAVGIGAVLFGIAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQORM 527
511 VOREKRAVGIGAVLFGIAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQOHL 570
528 LQLTWGIKQLOARLAVERYIGDOOLLGIWCGSKLICCTAVPNWASWSKSLDRINWN 587
571 LQLTWGIKQLOARLAVERYIGDOOLLGIWCGSKLICCTAVPNWASWSKSLDRINWN 630
588 MTWMEEREINDYTSEIYTLIEESONQOEKNEBELLEDKNASLNNWFDITWMLWY 643
631 MTWMEEREINDYTSEIYTLIEESONQOEKNEBELLEDKNASLNNWFDITWMLWY 686

```

Search completed: December 12, 2003, 12:34:26  
 Job time : 16.6215 secs

**THIS PAGE BLANK (US 710)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 26.8062 Seconds  
(without alignments)  
4461.192 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYYGVVPWKEATT.....ELDKMASIMWFDTITNLMWY 643

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 664280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 664280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10C\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3462	100.0	643	US-10-032-162-13	Sequence 13, App1
2	3295	95.2	625	US-10-032-162-17	Sequence 17, App1
3	3056	88.3	579	US-10-032-162-15	Sequence 15, App1
4	3041	87.8	842	US-10-190-435-2	Sequence 2, App1
5	3041	87.8	842	US-10-241-009-2	Sequence 2, App1
6	3041	87.8	842	US-10-190-434B-2	Sequence 2, App1
7	3041	87.8	842	US-10-190-305A-2	Sequence 2, App1
8	3041	87.8	842	US-09-476-242-2	Sequence 2, App1
9	2992.5	86.4	861	US-10-026-741-103	Sequence 103, App
10	2964	85.6	856	US-09-476-242-1	Sequence 1, App1
11	2953	85.3	856	US-10-196-515-11	Sequence 11, App1
12	2910.5	84.1	726	US-10-196-515-3	Sequence 3, App1
13	2897.5	83.7	759	US-10-196-515-12	Sequence 12, App1
14	2895.5	83.6	868	US-09-938-406-1	Sequence 1, App1
15	2862	82.7	619	US-09-891-609-4	Sequence 4, App1

16	2862	82.7	646	US-09-891-609-2	Sequence 2, App1
17	2736.5	79.0	860	US-10-190-435-6	Sequence 6, App1
18	2736.5	79.0	860	US-10-241-009-6	Sequence 6, App1
19	2736.5	79.0	860	US-10-190-434B-6	Sequence 6, App1
20	2736.5	79.0	860	US-10-190-305A-6	Sequence 6, App1
21	2696	77.9	867	US-10-190-435-3	Sequence 3, App1
22	2696	77.9	867	US-10-190-435-126	Sequence 126, App
23	2696	77.9	867	US-10-241-009-3	Sequence 3, App1
24	2696	77.9	867	US-10-190-434B-3	Sequence 3, App1
25	2696	77.9	867	US-10-190-305A-3	Sequence 3, App1
26	2694	77.8	858	US-10-190-435-150	Sequence 150, App
27	2687	77.6	869	US-10-190-435-4	Sequence 4, App1
28	2687	77.6	869	US-10-241-009-4	Sequence 4, App1
29	2687	77.6	869	US-10-190-434B-4	Sequence 4, App1
30	2687	77.6	869	US-10-190-305A-4	Sequence 4, App1
31	2684.5	77.5	803	US-10-190-435-135	Sequence 135, App
32	2683.5	77.5	803	US-10-190-435-134	Sequence 134, App
33	2683.5	77.5	870	US-10-190-435-127	Sequence 127, App
34	2682	77.5	861	US-10-190-435-139	Sequence 139, App
35	2670	77.1	870	US-10-190-435-147	Sequence 147, App
36	2669	77.1	862	US-10-190-435-141	Sequence 141, App
37	2668.5	77.1	865	US-10-190-435-140	Sequence 140, App
38	2668.5	77.0	857	US-10-190-435-138	Sequence 138, App
39	2666	77.0	849	US-10-190-435-148	Sequence 148, App
40	2662.5	76.9	855	US-10-190-435-144	Sequence 144, App
41	2660.5	76.8	845	US-10-190-435-129	Sequence 129, App
42	2660.5	76.8	845	US-10-190-435-130	Sequence 130, App
43	2648	76.5	855	US-10-369-294-11	Sequence 11, App1
44	2648	76.5	855	US-10-369-294-12	Sequence 12, App1
45	2644	76.4	844	US-09-991-258-19	Sequence 19, App1

ALIGNMENTS

RESULT 1:  
US-10-032-162-13  
; Sequence 13, Application US/10032162  
; Publication No. US2003052839A1  
GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELEKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADON J  
; APPLICANT: JOHN, MOORE P  
TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
FILE REFERENCE: 2048/59331az  
CURRENT APPLICATION NUMBER: US/10/032,162  
CURRENT FILING DATE: 2002-10-24  
PRIOR APPLICATION NUMBER: 09/602,864  
PRIOR FILING DATE: 2000-06-23  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
; LENGTH: 643  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-13  
Query Match 100.0%; Score 3462; DB 15; Length 643;  
Best Local Similarity 100.0%; Pred. No. 1.5e-308;  
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VEKLMVTYYGVVPWKEATTTLFCASDAKAYDTEVHNVMATHACVPTDPNQEVLLENT 60  
DB 1 VEKLMVTYYGVVPWKEATTTLFCASDAKAYDTEVHNVMATHACVPTDPNQEVLLENT 60  
QY EHFMMKNNNVEQOEEDIIISLMDOSLKPQVGLPTCLNCKQDVNATNTTDSGTERG 120  
DB EHFMMKNNNVEQOEEDIIISLMDOSLKPQVGLPTCLNCKQDVNATNTTDSGTERG 120  
QY EIKKSNITTSIRDEVOKEAYLFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180  
DB EIKKSNITTSIRDEVOKEAYLFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180

```

Db      121 EIRKCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180
Qy      181 IPIHYCAPAGFALLKNDKTFNGKGPCKNVSTVQCTHGIRPVYSTOLLNGSLAEERVI 240
Db      181 IPIHYCAPAGFALLKNDKTFNGKGPCKNVSTVQCTHGIRPVYSTOLLNGSLAEERVI 240
Qy      241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHANI 300
Db      241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHANI 300
Qy      301 SRKAMDTLKQIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQOLFNST 360
Db      301 SRKAMDTLKQIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQOLFNST 360
Qy      361 MNNTGSSNNTGNTTILPCRIKOIIMMOEYKAMAPPIRGQIRCSSNITGLLTRDG 420
Db      361 MNNTGSSNNTGNTTILPCRIKOIIMMOEYKAMAPPIRGQIRCSSNITGLLTRDG 420
Qy      421 GINENGTEIFRPGGDMRDMWRSEFYKYKVKIEPLGVAPTKCRRVQREKRAVGIGAV 480
Db      421 GINENGTEIFRPGGDMRDMWRSEFYKYKVKIEPLGVAPTKCRRVQREKRAVGIGAV 480
Qy      481 FLGFLGAAGSTMAASNTLTVQARLLSGIVQOONNLLRAIEAQORMLQLTWGIKQLOA 540
Db      481 FLGFLGAAGSTMAASNTLTVQARLLSGIVQOONNLLRAIEAQORMLQLTWGIKQLOA 540
Qy      541 RVLAVERYLDQDQLLGWGSGLICTTAVPMAASNSKSLDRIMNMNTWEMEREIDNY 600
Db      541 RVLAVERYLDQDQLLGWGSGLICTTAVPMAASNSKSLDRIMNMNTWEMEREIDNY 600
Qy      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLMY 643
Db      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLMY 643

```

```

RESULT 2
US-10-032-162-17
; Sequence 17, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331a2
; CURRENT APPLICATION NUMBER: US/10/032,162
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
US-10-032-162-17

```

Query Match 95.2%; Score 3295; DB 15; Length 625;  
Best Local Similarity 96.0%; Pred. No. 3.le-293;

```

Matches 617; Conservative 0; Mismatches 8; Indels 18; Gaps 1;
Qy      1 VEKLMATVYGVGVWMEATTTLCASDAXAYDEVYHNVATHACVPTDPPOQVLENT 60
Db      1 VEKLMATVYGVGVWMEATTTLCASDAXAYDEVYHNVATHACVPTDPPOQVLENT 60
Qy      61 EHNMMKNMNVQEQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERG 120
Db      61 EHNMMKNMNVQEQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERG 120
Qy      121 EIRKCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      121 EIRKCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180
Qy      181 IPIHYCAPAGFALLKNDKTFNGKGPCKNVSTVQCTHGIRPVYSTOLLNGSLAEERVI 240
Db      181 IPIHYCAPAGFALLKNDKTFNGKGPCKNVSTVQCTHGIRPVYSTOLLNGSLAEERVI 240
Qy      241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHANI 300
Db      241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHANI 300
Qy      301 SRKAMDTLKQIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQOLFNST 360
Db      301 SRKAMDTLKQIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQOLFNST 360
Qy      361 MNNTGSSNNTGNTTILPCRIKOIIMMOEYKAMAPPIRGQIRCSSNITGLLTRDG 420
Db      361 MNNTGSSNNTGNTTILPCRIKOIIMMOEYKAMAPPIRGQIRCSSNITGLLTRDG 420
Qy      421 GINENGTEIFRPGGDMRDMWRSEFYKYKVKIEPLGVAPTKCRRVQREKRAVGIGAV 480
Db      421 GINENGTEIFRPGGDMRDMWRSEFYKYKVKIEPLGVAPTKCRRVQREKRAVGIGAV 480
Qy      481 FLGFLGAAGSTMAASNTLTVQARLLSGIVQOONNLLRAIEAQORMLQLTWGIKQLOA 540
Db      481 FLGFLGAAGSTMAASNTLTVQARLLSGIVQOONNLLRAIEAQORMLQLTWGIKQLOA 540
Qy      541 RVLAVERYLDQDQLLGWGSGLICTTAVPMAASNSKSLDRIMNMNTWEMEREIDNY 600
Db      541 RVLAVERYLDQDQLLGWGSGLICTTAVPMAASNSKSLDRIMNMNTWEMEREIDNY 600
Qy      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLMY 643
Db      583 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLMY 625

```

```

RESULT 3
US-10-032-162-15
; Sequence 15, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331a2
; CURRENT APPLICATION NUMBER: US/10/032,162
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-15

```

Query Match 88.3%; Score 3056; DB 15; Length 579;  
Best Local Similarity 89.3%; Pred. No. 2.3e-271;

Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;

```

QY 1 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 60
QY 61 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 120
DB 61 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 120
QY 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 180
DB 99 -----GCDTSVITQACPKISFE 116
QY 181 IPIHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEVY 240
DB 117 IPIHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEVY 176
QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDIROAHCN 300
DB 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDIROAHCN 236
QY 301 SRAKNDTLKQIVIKLREQFENKTIIVFNHSSGDEPEIVMHSFNGCEFFYCNSTOLFNS 360
DB 237 SRAKNDTLKQIVIKLREQFENKTIIVFNHSSGDEPEIVMHSFNGCEFFYCNSTOLFNS 296
QY 361 WNNNTEGSSNTEGNTITLPCRIOIINMOEYKAMAPRIGQIRCSSNITGLLTRD 420
DB 297 WNNNTEGSSNTEGNTITLPCRIOIINMOEYKAMAPRIGQIRCSSNITGLLTRD 356
QY 421 GINENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGVAPTCKRVRVOREKRAVIGAV 480
DB 357 GINENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGVAPTCKRVRVOREKRAVIGAV 416
QY 481 FLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 540
DB 417 FLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 476
QY 541 RVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIMNNMTMMEWEREID 600
DB 477 RVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIMNNMTMMEWEREID 536
QY 601 TSEIYTLIEESONQOEKNEQELLEDKMASLNNMFDITNMLY 643
DB 537 TSEIYTLIEESONQOEKNEQELLEDKMASLNNMFDITNMLY 579

```

## RESULT 4

US-190-435-2  
Sequence 2, Application US/10190435  
Publication No. US20030143248A1

GENERAL INFORMATION:

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: BARNETT, Susan W.

APPLICANT: LIAN, Ying

APPLICANT: ENGELBRECHT, Susan

APPLICANT: VAN RENSBURG, Estrelita J.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

FILE REFERENCE: P18133.003 / 2302-18133

CURRENT APPLICATION NUMBER: US/10/190,435

NUMBER OF SEQ ID NOS: 319

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 842

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: SP162

US-10-190-435-2

Query Match

87.8%; Score 3041; DB 12; Length 842;

Best Local Similarity 88.7%; Pred. No. 9.6e-270; Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 60
QY 25 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 84
DB 25 VEKLMVTVYGVVPWKEATTTTLFCASDAKAYDTEVHNWMAHACVPTDPNQEVLLENT 84
QY 61 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 119
DB 61 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 119
QY 85 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 144
DB 85 EHFNMKNMWEQOMEDIISLMDQSLKPCVLTPLCTVLCNKQVATNTYNDSEGTMERG 144
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
DB 145 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEVY 239
DB 205 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEVY 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDIROAHCN 299
DB 265 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDIROAHCN 324
QY 300 ISRAKNDTLKQIVIKLREQFENKTIIVFNHSSGDEPEIVMHSFNGCEFFYCNSTOLFNS 359
DB 325 ISRAKNDTLKQIVIKLREQFENKTIIVFNHSSGDEPEIVMHSFNGCEFFYCNSTOLFNS 384
QY 360 TNNNTEGSSNTEGNTITLPCRIOIINMOEYKAMAPRIGQIRCSSNITGLLTRD 419
DB 385 TNNNTEGSSNTEGNTITLPCRIOIINMOEYKAMAPRIGQIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGVAPTCKRVRVOREKRAVIGAV 478
DB 443 GGIN-ENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGVAPTCKRVRVOREKRAVIGAV 502
QY 479 AVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 538
DB 503 AVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 562
QY 539 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIMNNMTMMEWEREID 598
DB 563 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIMNNMTMMEWEREID 622
QY 599 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNMFDITNMLY 643
DB 623 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNMFDITNMLY 667

```

## RESULT 5

US-10-241-009-2  
Sequence 2, Application US/10241009  
Publication No. US20030170614A1

GENERAL INFORMATION:

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: BARNETT, Susan

APPLICANT: LIAN, Ying

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

FILE REFERENCE: 2300-1621.21

CURRENT APPLICATION NUMBER: US/10/241,009

NUMBER OF SEQ ID NOS: 68

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 842

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: SP162

US-10-241-009-2

Query Match

Best Local Similarity 88.7%; Pred. No. 9.6e-270;

87.8%; Score 3041; DB 12; Length 842;

Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVVYGVPRVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 60
DB 25 VEKLMVTVVYGVPRVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 84
QY 61 EHFMMKNNWEOQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTNDSGEMER 119
DB 85 EHFMMKNNWEOQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTNDSGEMER 144
QY 120 GEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRILSCDTSVITQACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNMQKEYALFYKLDVVPIDNNNTSYKLINCNTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALCKNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGSLAEGBV 239
DB 205 PIPHYCAPAGFALCKNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGSLAEGBV 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 299
DB 265 IRSENFTDNKTIIVOLKESVEINCTRPNNNTRKSIITIGRAFYATGDIIGDIRQAHN 324
QY 300 ISRAKMDTLKQIVIKRBEQENKTIYFNHSSGSDPEIWNHSEFCGEGFFYCNSDTOLFNS 359
DB 325 ISGKMNNTLKQIVTKLQAFGNKTIYFKOSSGSDPEIWNHSEFCGEGFFYCNSDTOLFNS 384
QY 360 TANNTEGSNNTEGNTTLPCRIRKQIINMWOEVKAMVAPIRGOIRCSSNITGLLTRD 419
DB 385 TW-NNTIGPNNTNG-TITLPCRIRKQIINMWOEVKAMVAPIRGOIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRPGGDMRDNMRSEFYKYKVKIEPLGVAFTKCKRRVVQREKRAVIG 478
DB 443 GKKEISNTTEIFRPGGDMRDNMRSELYKYKVKIEPLGVAFTKCKRRVVQREKRAVIG 502
QY 479 AVPLIGFGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQMLQITWGIKOL 538
DB 503 AMPLIGFGAAGSTMGASMTLTVQARQLLSGIVQOQNNLLRAIEAOQMLQITWGIKOL 562
QY 539 QARLAVERYLKQOQLGIMCGSKLICTTAVPNWASMSKSLDRIMNNMTWMEWERID 598
DB 563 QARLAVERYLKQOQLGIMCGSKLICTTAVPNWASMSKSLDRIMNNMTWMEWERID 622
QY 599 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWEDITNMLWY 643
DB 623 NYTSLIYTLIEESQNOQEKNEQELLEDKMASLMMWEDISKMLWY 667

```

```

RESULT 6
US-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGERDE, Jan
; APPLICANT: BARNETT, Susan
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/10/190.434B
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-434B-2

```

Query Match 87.8%; Score 3041; DB 12; Length 842;  
 Best Local Similarity 88.7%; Pred. No. 9.6e-270;  
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVVYGVPRVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 60
DB 25 VEKLMVTVVYGVPRVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 84
QY 61 EHFMMKNNWEOQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTNDSGEMER 119
DB 85 EHFMMKNNWEOQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTNDSGEMER 144
QY 120 GEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRILSCDTSVITQACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNMQKEYALFYKLDVVPIDNNNTSYKLINCNTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALCKNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGSLAEGBV 239
DB 205 PIPHYCAPAGFALCKNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGSLAEGBV 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 299
DB 265 IRSENFTDNKTIIVOLKESVEINCTRPNNNTRKSIITIGRAFYATGDIIGDIRQAHN 324
QY 300 ISRAKMDTLKQIVIKRBEQENKTIYFNHSSGSDPEIWNHSEFCGEGFFYCNSDTOLFNS 359
DB 325 ISGKMNNTLKQIVTKLQAFGNKTIYFKOSSGSDPEIWNHSEFCGEGFFYCNSDTOLFNS 384
QY 360 TANNTEGSNNTEGNTTLPCRIRKQIINMWOEVKAMVAPIRGOIRCSSNITGLLTRD 419
DB 385 TW-NNTIGPNNTNG-TITLPCRIRKQIINMWOEVKAMVAPIRGOIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRPGGDMRDNMRSEFYKYKVKIEPLGVAFTKCKRRVVQREKRAVIG 478
DB 443 GKKEISNTTEIFRPGGDMRDNMRSELYKYKVKIEPLGVAFTKCKRRVVQREKRAVIG 502
QY 479 AVPLIGFGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQMLQITWGIKOL 538
DB 503 AMPLIGFGAAGSTMGASMTLTVQARQLLSGIVQOQNNLLRAIEAOQMLQITWGIKOL 562
QY 539 QARLAVERYLKQOQLGIMCGSKLICTTAVPNWASMSKSLDRIMNNMTWMEWERID 598
DB 563 QARLAVERYLKQOQLGIMCGSKLICTTAVPNWASMSKSLDRIMNNMTWMEWERID 622
QY 599 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWEDITNMLWY 643
DB 623 NYTSLIYTLIEESQNOQEKNEQELLEDKMASLMMWEDISKMLWY 667

```

```

RESULT 7
US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGERDE, Jan
; APPLICANT: BARNETT, Susan
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190.305A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-305A-2

```

Query Match 87.8%; Score 3041; DB 12; Length 842;  
 Best Local Similarity 88.7%; Pred. No. 9.6e-270;  
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

Qy 1 VEKLVVTVYGVVWKEATTTLCASDAKAYDEVHNVATHACVPTDNPQEVILENT 60
Db 25 VEKLVVTVYGVVWKEATTTLCASDAKAYDEVHNVATHACVPTDNPQEVILENT 84
Qy 61 EHFNNKNNMVEQOMEDIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTDSEGITER 119
Db 85 EHFNNKNNMVEQOMEDIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTDSEGITER 144
Qy 120 GEIKNCSNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 145 GEIKNCSKVTTSIRNKQKEVALFYKLDVPIIDNDNTSYKLNCTNSVITQACPKVSFE 204
Qy 180 PIPHYCAPAGFAILKCDKTFNGKPCKNVSTVOCTGHGIRPVSTOLLNGLSIAEEVY 239
Db 205 PIPHYCAPAGFAILKCDKTFNGKPCKNVSTVOCTGHGIRPVSTOLLNGLSIAEEVY 264
Qy 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTKRSIHGPGAFYTTGIIIGDIOAHCN 299
Db 265 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTKRSIHGPGAFYTTGIIIGDIOAHCN 324
Qy 300 ISRAKMDTLKQIVYKLEBOENKTIIVFNHSSGDPBEIVMHSFNGEGEFPYCNSDTQFN 359
Db 325 ISGEKMNNTLQIVYKLEBOENKTIIVFNHSSGDPBEIVMHSFNGEGEFPYCNSDTQFN 384
Qy 360 TNNNTGSSNTEGNTITLPCRIOIIMMOEVEGKAMTAPPIRGQIRCSSNITGLLTRD 419
Db 385 TNNNTGSSNTEGNTITLPCRIOIIMMOEVEGKAMTAPPIRGQIRCSSNITGLLTRD 442
Qy 420 GGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 478
Db 443 GKEISNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 502
Qy 479 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 538
Db 503 AMFLGFLGAGSTMGARSLTLTVQARQLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 562
Qy 539 QARVLAVERYLKDOQLIGWCSGKLICTTAVPNNASNSKSLDRIINNMTMWEREID 598
Db 563 QARVLAVERYLKDOQLIGWCSGKLICTTAVPNNASNSKSLDRIINNMTMWEREID 622
Qy 599 NYTSEIYTLIESONQOEKNEQELLELDKMSLNMWPDITWMLWY 643
Db 623 NYTNIYTLIESONQOEKNEQELLELDKMSLNMWPDITWMLWY 667

```

RESULT 8  
US-09-476-242-2  
Sequence 2, Application US/09476242  
Patent No. US20020146683A1

GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: HARTOG, Karin  
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
FILE REFERENCE: 1605.002  
CURRENT APPLICATION NUMBER: US/09/476,242  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-476-242-2

Query Match 87.8%; Score 3041; DB 10; Length 847;  
Best Local Similarity 88.7%; Pred. No. 9.7e-270;  
Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

Qy 1 VEKLVVTVYGVVWKEATTTLCASDAKAYDEVHNVATHACVPTDNPQEVILENT 60  
Db 30 VEKLVVTVYGVVWKEATTTLCASDAKAYDEVHNVATHACVPTDNPQEVILENT 89

```

Qy 61 EHFNNKNNMVEQOMEDIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTDSEGITER 119
Db 90 EHFNNKNNMVEQOMEDIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTDSEGITER 149
Qy 120 GEIKNCSNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 150 GEIKNCSKVTTSIRNKQKEVALFYKLDVPIIDNDNTSYKLNCTNSVITQACPKVSFE 209
Qy 180 PIPHYCAPAGFAILKCDKTFNGKPCKNVSTVOCTGHGIRPVSTOLLNGLSIAEEVY 239
Db 210 PIPHYCAPAGFAILKCDKTFNGKPCKNVSTVOCTGHGIRPVSTOLLNGLSIAEEVY 269
Qy 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTKRSIHGPGAFYTTGIIIGDIOAHCN 299
Db 270 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTKRSIHGPGAFYTTGIIIGDIOAHCN 329
Qy 300 ISRAKMDTLKQIVYKLEBOENKTIIVFNHSSGDPBEIVMHSFNGEGEFPYCNSDTQFN 359
Db 330 ISGEKMNNTLQIVYKLEBOENKTIIVFNHSSGDPBEIVMHSFNGEGEFPYCNSDTQFN 389
Qy 360 TNNNTGSSNTEGNTITLPCRIOIIMMOEVEGKAMTAPPIRGQIRCSSNITGLLTRD 419
Db 390 TNNNTGSSNTEGNTITLPCRIOIIMMOEVEGKAMTAPPIRGQIRCSSNITGLLTRD 447
Qy 420 GGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 478
Db 448 GKEISNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 507
Qy 479 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 538
Db 508 AMFLGFLGAGSTMGARSLTLTVQARQLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 567
Qy 539 QARVLAVERYLKDOQLIGWCSGKLICTTAVPNNASNSKSLDRIINNMTMWEREID 598
Db 568 QARVLAVERYLKDOQLIGWCSGKLICTTAVPNNASNSKSLDRIINNMTMWEREID 627
Qy 599 NYTSEIYTLIESONQOEKNEQELLELDKMSLNMWPDITWMLWY 643
Db 628 NYTNIYTLIESONQOEKNEQELLELDKMSLNMWPDITWMLWY 672

```

RESULT 9  
US-10-026-741-103  
Sequence 103, Application US/10026741  
Publication No. US20030049604A1

GENERAL INFORMATION:  
APPLICANT: CHARNEAU, PIERRE  
CLAVEL, FRANCOISE  
BORMAN, ANDREW  
OUILLENT, CAROLINE  
GUETARD, DENISE  
MONTAGNIER, LUC  
DONJON DE SAINT-MARTIN, JACQUELINE  
COHEN, JACQUES  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
SUBTYPE) ANTIGENS

NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &  
Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/026,741  
FILING DATE: 27-Dec-2001

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/817,441  
 FILING DATE: 31-AUG-1998  
 APPLICATION NUMBER: PCT/FR 95/01391  
 FILING DATE: 20-OCT-1995  
 APPLICATION NUMBER: FR 9412554  
 FILING DATE: 20-OCT-1994  
 APPLICATION NUMBER: FR 9502526  
 FILING DATE: 03-MAR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meyers, Kenneth J.  
 REGISTRATION NUMBER: 25,146  
 REFERENCE/DOCKET NUMBER: 03260.6005-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4400  
 INFORMATION FOR SEQ ID NO: 103:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
 US-10-026-741-103

Query Match 86.4%; Score 2992.5; DB 15; Length 861;  
 Best Local Similarity 86.7%; Pred. No. 2.8e-265; Indels 15; Gaps 8;  
 Matches 569; Conservative 32; Mismatches 40;

2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHNMWATHACVPTDPNPOEVLEAVTE 61  
 32 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHNMWATHACVPTDPNPOEVLEAVTE 91  
 62 HFNWKNMWEQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKOV-NAYNT---TNDSEG- 115  
 92 NFNWKNMWEQWQEDIIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTSSNTSSSGE 151  
 116 -TMRGEIKNGCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSVRLISCDSVITQACP 174  
 152 MMEKGEIKNGCSFNITTSIRDEVQKEYALFYKLDIIPIDNDTSTYKLSCTNSVTQACP 211  
 175 KISFEPPIHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVYSTOLLNGSLA 234  
 212 KVSFEPPIHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVYSTOLLNGSLA 271  
 235 EEEVVISRANFTDNAKTIIVQLNOSVEINCTRPNNNTKRSIRI-GRGPAFTTGEIIGD 292  
 272 EEEVVISRANFTDNAKTIIVQLNOSVEINCTRPNNNTKRSIRI-GRGPAFTTGEIIGD 330  
 293 IROAHCHNSRAKMDTLKQIVIKLREOF-ENKTIYFNHSSGGDEIIVHSGCEGFYFC 351  
 331 MROAHCHNSRAKMDTLKQIVIKLREOF-ENKTIYFNHSSGGDEIIVHSGCEGFYFC 390  
 352 NSTOLFSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMVAPPIRGQIR 407  
 391 NSTOLFSTWNNSTWSTGSGNTEGSDITTLPCRILKQIINMOWEYKAMVAPPIRGQIR 450  
 408 SSNTITGLLTRDGINENGTETIFRPGGDMRDNRSEFYKYKVIKIEPLGVAFTKCKRRV 467  
 451 SSNTITGLLTRDGINENGTETIFRPGGDMRDNRSEFYKYKVIKIEPLGVAFTKCKRRV 510  
 468 VQREKRAVIGLAVLGLFGAAGSTMGASMTLTVQARLLLSGIYVOOQNLLRAIEAOORM 527  
 511 VQREKRAVIGLAVLGLFGAAGSTMGASMTLTVQARLLLSGIYVOOQNLLRAIEAOORM 570  
 528 LQITVWGKIQOLARILAVERYLKDQQLIGWCSGSKLICTTAVPNNASMSKSLDRINWNN 587  
 571 LQITVWGKIQOLARILAVERYLKDQQLIGWCSGSKLICTTAVPNNASMSKSLDRINWNN 630  
 588 MTWMEWEIDNYTSEIYTLIEESQNOQEKNOBELLELDKMASLWNPEDITWMLY 643  
 631 MTWMEWEIDNYTSEIYTLIEESQNOQEKNOBELLELDKMASLWNPEDITWMLY 686

## RESULT 10

US-09-476-242-1  
 Sequence 1, Application US/09476242  
 Patent No. US2002014663A1  
 GENERAL INFORMATION:  
 APPLICANT: BARRETT, Susan  
 APPLICANT: HARTOG, Karin  
 APPLICANT: MARTIN, Eric  
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
 FILE REFERENCE: 1605.002  
 CURRENT APPLICATION NUMBER: US/09/476,242  
 CURRENT FILING DATE: 1999-12-30  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1  
 LENGTH: 856  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus  
 US-09-476-242-1

Query Match 85.6%; Score 2964; DB 10; Length 856;  
 Best Local Similarity 86.0%; Pred. No. 1.1e-262; Indels 10; Gaps 6;  
 Matches 560; Conservative 36; Mismatches 45;

2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHNMWATHACVPTDPNPOEVLEAVTE 61  
 32 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHNMWATHACVPTDPNPOEVLEAVTE 91  
 62 HFNWKNMWEQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKOV-NAYNT---TNDSEG- 119  
 92 NFNWKNMWEQWQEDIIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTSSNTSSSGE 151  
 120 GEIKNGCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSVRLISCDSVITQACP 179  
 152 GEIKNGCSFNITTSIRDEVQKEYALFYKLDIIPIDNDTSTYKLSCTNSVTQACPVSFE 211  
 180 PIPHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVYSTOLLNGSLAEEV 239  
 212 PIPHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVYSTOLLNGSLAEEV 271  
 240 IRSDFNTNNAKTIIVQLNOSVEINCTRPNNNTKRSIRI-GRGPAFTTGEIIGDIRQAH 297  
 272 IRSDFNTNNAKTIIVQLNOSVEINCTRPNNNTKRSIRI-GRGPAFTTGEIIGDIRQAH 330  
 298 CNISRAKMDTLKQIVIKLREOF-ENKTIYFNHSSGGDEIIVHSGCEGFYFCNSTOL 356  
 331 CNISRAKMDTLKQIVIKLREOF-ENKTIYFNHSSGGDEIIVHSGCEGFYFCNSTOL 390  
 357 FNSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMVAPPIRGQIRCSNIT 412  
 391 FNSTWNNSTWSTGSGNTEGSDITTLPCRILKQIINMOWEYKAMVAPPIRGQIRCSNIT 450  
 413 GLLTRDGINENGTETIFRPGGDMRDNRSEFYKYKVIKIEPLGVAFTKCKRRVOREK 472  
 451 GLLTRDGINENGTETIFRPGGDMRDNRSEFYKYKVIKIEPLGVAFTKCKRRVOREK 510  
 473 RAVGIGAVLGLFGAAGSTMGASMTLTVQARLLLSGIYVOOQNLLRAIEAOORMDLTV 532  
 511 RAVGIGAVLGLFGAAGSTMGASMTLTVQARLLLSGIYVOOQNLLRAIEAOORMDLTV 570  
 533 WGIQOLARILAVERYLKDQQLIGWCSGSKLICTTAVPNNASMSKSLDRINWNNWME 592  
 571 WGIQOLARILAVERYLKDQQLIGWCSGSKLICTTAVPNNASMSKSLDRINWNNWME 630  
 593 WEREIDNYTSEIYTLIEESQNOQEKNOBELLELDKMASLWNPEDITWMLY 643  
 631 WEREIDNYTSEIYTLIEESQNOQEKNOBELLELDKMASLWNPEDITWMLY 681

RESULT 11  
 US-10-196-515-11

```

; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: Hoxie, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

```

```

Query Match      85.3%; Score 2953; DB 15; Length 856;
Best Local Similarity 85.9%; Pred. No. 1.2e-261;
Matches 559; Conservative 36; Mismatches 46; Indels 10; Gaps 6;

QY 2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVVLENTTE 61
DB 32 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVVLENTTE 91
QY 62 HFNMMKNMVEQMEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 92 NFDMMKNDMEQMEDIIISLDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSGGMINEK 151
QY 120 GEIKNCSFNITTSIRDEYQKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEYQKEYALFYKLDIIPIDNTTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAFALIKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAEEV 239
DB 212 PIPHYCAPAFALIKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAEEV 271
QY 240 IRSDFNNAKTIIVOLKESYEINCTRPNNNRKSIH--GPRGAFYTTGEIIGDIRQAH 297
DB 272 IRSVNFIDNAKTIIVOLNTSVSEINCTRPNNNRKRIIRIGRGAFAFVIGK-IGNMROAH 330
QY 298 CNISRAKKNNDTLKQIVIKLREOF-ENKTIIVFNHSSGDBPEIVMHSFNCEGEFFYCNSTOL 356
DB 331 CNISRAKKNNDTLKQIVIKLREOF-ENKTIIVFNHSSGDBPEIVMHSFNCEGEFFYCNSTOL 390
QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRKIOIIMMOEYKAMAYAPPIRQOICSSNITL 412
DB 391 FNSTNN--NTEGSNNTEG-NTITLPCRKIOIIMMOEYKAMAYAPPIRQOICSSNITL 450
QY 413 GLLTRDGINENGEIIFRPGGDMRDNMRSSEFYKYVVKIEPLGVAPTKCRVRVOREK 472
DB 451 GLLTRDGINENGEIIFRPGGDMRDNMRSSEFYKYVVKIEPLGVAPTKCRVRVOREK 510
QY 473 RAVGIGAVFLGFGAAGSTMGASMTLVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 532
DB 511 RAVGIGAVFLGFGAAGSTMGASMTLVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 570
QY 533 WGIQOLARVLAVERYLGDQOLLGIMGCSGKLICTTAVPNNASNNKSLDRINMMNTME 592
DB 571 WGIQOLARVLAVERYLGDQOLLGIMGCSGKLICTTAVPNNASNNKSLDRINMMNTME 630
QY 593 WEREIDNTSEIYTLIEESQNOEQNEOELELDKMASLMMNFDITNMLWY 643
DB 631 WEREIDNTSEIYTLIEESQNOEQNEOELELDKMASLMMNFDITNMLWY 681

```

```

RESULT 12
US-10-196-515-3
; Sequence 3, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: Hoxie, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

```

```

Query Match      84.1%; Score 2910.5; DB 15; Length 726;
Best Local Similarity 84.7%; Pred. No. 7.2e-258;
Matches 549; Conservative 39; Mismatches 51; Indels 9; Gaps 6;

QY 2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVVLENTTE 61
DB 32 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVVLENTTE 91
QY 62 HFNMMKNMVEQMEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 92 NFNMMKNDMEQMEDIIISLDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSGGMINEK 151
QY 120 GEIKNCSFNITTSIRDEYQKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEYQKEYALFYKLDIIPIDNTTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAFALIKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAEEV 239
DB 212 PIPHYCAPAFALIKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAEEV 271
QY 240 IRSDFNNAKTIIVOLKESYEINCTRPNNNRK--SIHIGRGAFAFVTTGEIIGDIRQAH 297
DB 272 IRSVNFIDNAKTIIVOLNTSVSEINCTRPNNNRKRIIRIGRGAFAFVIGK-IGNMROAH 330
QY 298 CNISRAKKNNDTLKQIVIKLREOF-ENKTIIVFNHSSGDBPEIVMHSFNCEGEFFYCNSTOL 356
DB 331 CNISRAKKNNDTLKQIVIKLREOF-ENKTIIVFNHSSGDBPEIVMHSFNCEGEFFYCNSTOL 390
QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRKIOIIMMOEYKAMAYAPPIRQOICSSNITL 415
DB 391 FNSTNN--NTEGSNNTEG-NTITLPCRKIOIIMMOEYKAMAYAPPIRQOICSSNITL 448
QY 416 LTRDGINENGEIIFRPGGDMRDNMRSSEFYKYVVKIEPLGVAPTKCRVRVOREK 475
DB 449 LTRDGINENGEIIFRPGGDMRDNMRSSEFYKYVVKIEPLGVAPTKCRVRVOREK 508
QY 476 GIGAVFLGFGAAGSTMGASMTLVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 535
DB 509 GIGAVFLGFGAAGSTMGASMTLVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 568
QY 536 KQOLARVLAVERYLGDQOLLGIMGCSGKLICTTAVPNNASNNKSLDRINMMNTME 595
DB 569 KQOLARVLAVERYLGDQOLLGIMGCSGKLICTTAVPNNASNNKSLDRINMMNTME 628
QY 596 EIDNTSEIYTLIEESQNOEQNEOELELDKMASLMMNFDITNMLWY 643
DB 629 EIDNTSEIYTLIEESQNOEQNEOELELDKMASLMMNFDITNMLWY 676

```

RESULT 13  
US-10-196-515-12  
; Sequence 12, Application US/10196515  
; Publication No. US20030091594A1  
; GENERAL INFORMATION:  
; APPLICANT: HOXIE, James A.  
; APPLICANT: LABRANCHE, Celia C.  
; APPLICANT: DOMS, Robert W.  
; APPLICANT: HOFEMAN, Trevor L.  
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
; FILE REFERENCE: Hoxie 9596-104U1 (0282)  
; CURRENT APPLICATION NUMBER: US/10/196,515  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: US/09/337,387  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: US 09/317,556  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 759  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-12

Query Match 83.7%; Score 2897.5; DB 15; Length 759;  
Best Local Similarity 84.6%; Pred. No. 1,2e-256;  
Matches 548; Conservative 40; Mismatches 51; Indels 9; Gaps 6;  
QY 2 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATACVPTDNPQEVLENTVE 61  
DB 32 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATACVPTDNPQEVLENTVE 91  
QY 62 HFNWKNWVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 119  
DB 92 NFNWKNWVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSGSRMMEK 151  
QY 120 GEIKNSFNITTSIRDEQKRYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPISFE 179  
DB 152 GEIKNSFNITTSIRDEQKRYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPISFE 211  
QY 180 PIRHYCAPGFALCKNDKTFNGKPCKNVSTVOCTHGRPVYSTOLLNGSLAESEV 239  
DB 212 PIRHYCAPGFALCKNDKTFNGKPCKNVSTVOCTHGRPVYSTOLLNGSLAESEV 271  
QY 240 IRSDFNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIGDIRQAH 297  
DB 272 IRSDFNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIGDIRQAH 330  
QY 298 CNISRAKMDTLKOIVIKLREOF--ENKTIYFNHSSGGDPEIVMHSFNGEGEFPYCSTOL 356  
DB 331 CNISRAKMDTLKOIVIKLREOF--ENKTIYFNHSSGGDPEIVMHSFNGEGEFPYCSTOL 390  
QY 357 FNSWNNNTSNTTGG--NTTLPCRIKQIINNMQEVKAMYPAPINGOIRCSNITGL 415  
DB 391 FNSWNNNTSNTTGG--NTTLPCRIKQIINNMQEVKAMYPAPINGOIRCSNITGL 448  
QY 416 LTRGGINENGTETFRPGGDMRDNRSEFYKVKVLEPLGVAPTCKRRAVVRKRAV 475  
DB 449 LTRGGINENGTETFRPGGDMRDNRSEFYKVKVLEPLGVAPTCKRRAVVRKRAV 508  
QY 476 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQEMQLTWGI 535  
DB 509 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQEMQLTWGI 568  
QY 536 KQLOARVLAVERVYIGDOQLLGWCSGKLICTTAVPWNASWSNLSLRINNMWMEWER 595  
DB 569 KQLOARVLAVERVYIGDOQLLGWCSGKLICTTAVPWNASWSNLSLRINNMWMEWER 628  
QY 596 EIDNYTSEIVTLIESQNOQEKNEQELLEDKMASLWNVFDTIMWLV 643

DB 629 EINNYSILHSLIESQNOQEKNEQELLEDKMASLWNVFDTIMWLV 676  
RESULT 14  
US-09-938-406-1  
; Sequence 1, Application US/09938406  
; Patent No. US20020155120A1  
; GENERAL INFORMATION:  
; APPLICANT: Lowell, George  
; APPLICANT: Vancolt, Thomas  
; APPLICANT: Birt, Deborah  
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR  
; FILE REFERENCE: 40646-20002.10  
; CURRENT APPLICATION NUMBER: US/09/938,406  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US 97/12253  
; PRIOR FILING DATE: 1997-07-10  
; PRIOR APPLICATION NUMBER: US 60/021,687  
; PRIOR FILING DATE: 1996-07-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Virus HIV-1  
US-09-938-406-1

Query Match 83.6%; Score 2895.5; DB 10; Length 868;  
Best Local Similarity 82.1%; Pred. No. 2,2e-256;  
Matches 541; Conservative 40; Mismatches 59; Indels 19; Gaps 4;  
QY 4 LMTVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATACVPTDNPQEVLENTVE 63  
DB 35 LMTVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATACVPTDNPQEVLENTVE 94  
QY 64 NFNWKNWVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSGTM----- 117  
DB 95 NFNWKNWVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSGTM----- 154  
QY 118 -BERGEIKNSFNITTSIRDEQKRYALFYKLDVVPIDN-----NNTSYLLISCDTSVITQ 171  
DB 155 RKGEMNCSFNITTSIRDEQKRYALFYKLDVVPIDN-----NNTSYLLISCDTSVITQ 214  
QY 172 ACPKISFEPIPIHYCAPGFALCKNDKTFNGKPCKNVSTVOCTHGRPVYSTOLLNG 231  
DB 215 ACPKISFEPIPIHYCAPGFALCKNDKTFNGKPCKNVSTVOCTHGRPVYSTOLLNG 274  
QY 232 SLAESEVYISDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIG 291  
DB 275 SLAESEVYISDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIG 334  
QY 292 DTRQAHNISRANKMDTLKOIVIKLREOFENKTIYFNHSSGGDPEIVMHSFNGEGEFPYC 351  
DB 335 DTRQAHNISRANKMDTLKOIVIKLREOFENKTIYFNHSSGGDPEIVMHSFNGEGEFPYC 394  
QY 352 NSTOLFNSWNNNTSNTTGG-----NTEGNTTLPCRIKQIINNMQEVKAMYPAPINGOIR 405  
DB 395 NSTOLFNSWNNNTSNTTGG-----NTEGNTTLPCRIKQIINNMQEVKAMYPAPINGOIR 454  
QY 406 RGSNITGLLITRDGGINENGTETFRPGGDMRDNRSEFYKVKVLEPLGVAPTCKR 465  
DB 445 RGSNITGLLITRDGGINENGTETFRPGGDMRDNRSEFYKVKVLEPLGVAPTCKR 514  
QY 466 RYVQREKRAVG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQ 524  
DB 515 RYVQREKRAVG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQ 574  
QY 525 QRMQLTWGIIKQLOARVLAVERVYIGDOQLLGWCSGKLICTTAVPWNASWSNLSLRIN 584  
DB 574 QRMQLTWGIIKQLOARVLAVERVYIGDOQLLGWCSGKLICTTAVPWNASWSNLSLRIN 628



Db 575 OHLLQLTWGIKQJLQARILAVERYKDDQQLGFWGCSGLICTTAVPNNASNSKTLDOI 634  
 QY 585 WNNMTWMEWEREIDNYTSEIYTLIESONQOEKNEQELLELDKWSLNNWFDITNMLWY 643  
 Db 635 WNNMTWMEWEREIDNYTHLYTLIESONQOEKNEQELLELDKWSLNNWFDITNMLWY 693

Db 573 NYTNLIYTLIESONQOEKNEQELLELDKWSLNNWFDISKWLY 617  
 Search completed: December 12, 2003, 12:45:21  
 Job time : 28.8062 secs

RESULT 15  
 US-09-891-609-4  
 ; Sequence 4, Application US/09891609  
 ; Patent No. US20020127238A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stamatatos, Leonidas  
 ; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
 ; FILE REFERENCE: 2570-1-002N  
 ; CURRENT APPLICATION NUMBER: US/09/891,609  
 ; PRIOR FILING DATE: 2001-06-26  
 ; PRIOR FILING DATE: 2000-06-27  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 619  
 ; TYPE: PRT  
 ; ORGANISM: Human immunodeficiency virus type 1  
 US-09-891-609-4

Query Match 82.7%; Score 2862; DB 10; Length 619;

Best local similarity 84.5%; Pred. No. 1.6e-253;

Matches 545; Conservative 25; Mismatches 43; Indels 32; Gaps 6;

QY 1 VEKLVATVYVGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60  
 Db 3 VEKLVATVYVGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 62  
 QY 61 EHPNMMKNNWVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCDV-NATNTNDSGTWER 119  
 Db 63 ENFNMMKNNWVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCDV-NATNTNDSGTWER 122  
 QY 120 GEIKKCSFNITTSIDDEVQKEVALFYKLDVPRIDNNNTSYRLISCDTSVITQACPKISFE 179  
 Db 123 GEIKKCSFVVGAG-----KLINCNTSVITQACPKVSFE 155  
 QY 180 PIPHYCAPAGFAILKCDNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLAEBEV 239  
 Db 156 PIPHYCAPAGFAILKCDNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLAEBEV 215  
 QY 240 IRSDFNFTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRQAHCN 299  
 Db 216 IRSENFNTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYATGDIIGDIRQAHCN 275  
 QY 300 ISRAKMNNTLKOIYVLAKEQFENKTIIVFNHSGGDEIYHMSFNCGEFFYCNSTOLFNS 359  
 Db 276 ISGERMNNTLKOIYVLAKEQFENKTIIVFNHSGGDEIYHMSFNCGEFFYCNSTOLFNS 335  
 QY 360 TWNNNTGSSNNTGNTTLLPCRIKOIINMMOEVGKAMTAPPIRGQIRCSSNITGLLTRD 419  
 Db 336 TW-NNTIGPNNNTG-TTLLPCRIKOIINMMOEVGKAMTAPPIRGQIRCSSNITGLLTRD 393  
 QY 420 GGIN-ENGTEIFRPGGDMRDWRSEFYKYKVKIEPLGVAFTCKRRVQREKRAVGIG 478  
 Db 394 GGKEISNTTEIFRPGGDMRDWRSELYKYKVKIEPLGVAFTKARRVQREKRAVTLG 453  
 QY 479 AVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQLTWGIKOL 538  
 Db 454 AMFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQLTWGIK-L 512  
 QY 539 QARVLAVERYLGDQQLGFWGCSGLICTTAVPNNASNSKSLDRIMNNMTWMEWEREID 598  
 Db 513 QARVLAVERYLKDQQLGFWGCSGLICTTAVPNNASNSKSLDOIWNNMTWMEWEREID 572  
 QY 599 NTSSEIYTLIESONQOEKNEQELLELDKWSLNNWFDITNMLWY 643

**THIS PAGE BLANK (US 70)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:27:29 ; Search time 14.9697 Seconds

(without alignments)  
4130.780 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYYGVVPWKREATT.....ELDKWASIMNWDITNLMWY 643

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 283308 segs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3424	98.9	847	2	envelope glycoprot
2	3395	98.1	847	2	env protein - huma
3	3007.5	86.9	852	2	envelope glycoprot
4	3000	86.7	843	1	env polypotein pr
5	2992.5	86.4	861	1	env polypotein pr
6	2989	86.3	852	1	env polypotein -
7	2978	86.0	856	1	env polypotein pr
8	2974	85.9	854	1	env polypotein pr
9	2972	85.8	854	2	env protein - huma
10	2972	85.8	855	1	env polypotein pr
11	2967.5	85.7	859	1	env polypotein pr
12	2967	85.7	861	1	env polypotein pr
13	2958.5	85.5	856	1	env polypotein pr
14	2953.5	85.3	851	2	env polypotein pr
15	2895.5	83.6	729	1	env polypotein pr
16	2895.5	83.6	861	1	env polypotein pr
17	2895.5	83.6	868	1	env polypotein pr
18	2706	78.2	855	1	env polypotein pr
19	2686	77.6	853	2	envelope glycoprote
20	2657.5	76.8	859	2	envelope glycoprote
21	2638.5	76.2	856	1	env polypotein pr
22	2614.5	75.5	846	1	env polypotein pr
23	2244	64.8	854	2	env polypotein pr
24	2165	62.5	506	2	env polypotein pr
25	2002.5	57.8	443	2	env polypotein pr
26	1985.5	57.4	445	2	env polypotein pr
27	1938.5	56.0	495	2	env polypotein pr
28	1864	53.8	877	2	envelope protein p
29	1846	53.3	454	2	env polypotein p

30	1825	52.7	863	2	gag polypotein -
31	1320.5	38.1	290	2	env protein - huma
32	1299.5	37.5	297	2	envelope glycoprote
33	1206	34.8	859	1	env polypotein pr
34	1204.5	34.8	852	1	env polypotein pr
35	1201.5	34.7	299	2	envelope glycoprote
36	1196.5	34.6	299	2	envelope glycoprote
37	1191.5	34.4	299	2	envelope glycoprote
38	1188	34.3	712	1	env polypotein pr
39	1187.5	34.3	299	2	envelope glycoprote
40	1186.5	34.3	299	2	envelope glycoprote
41	1186	34.3	294	2	envelope glycoprote
42	1184	34.2	294	2	envelope glycoprote
43	1183.5	34.2	299	2	envelope glycoprote
44	1180.5	34.1	299	2	envelope glycoprote
45	1180	34.1	877	2	env polypotein -

ALIGNMENTS

RESULT 1									
T09448									
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRL)									
C:Species: human immunodeficiency virus type 1, HIV-1									
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999									
C:Accession: T09448									
R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie, submitted to the EMBL Data Library, July 1996									
A:Reference number: Z16673									
A:Status: preliminary; translated from GB/EMBL/DBJ									
A:Molecule type: DNA									
A:Residues: 1-847 <PAM>									
A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781									
C:Genetics:									
A:Gene: env									
C:Superfamily: type B retrovirus env polypotein									
Query Match									
Best Local Similarity 98.9%; Score 3424; DB 2; Length 847;									
Matches 638; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	VEKLMVTYYGVVPWKREATTLP	CASDAKADTEVHNWATHACVDPDPNPOEVLENT	60					
Db	30	VEKLMVTYYGVVPWKREATTLP	CASDAKADTEVHNWATHACVDPDPNPOEVLENT	89					
Qy	61	EHFMWKNWVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTNDSGTMERG	120						
Db	90	EHFMWKNWVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTNDSGTMERG	149						
Qy	121	EIKKCSFRTTTSIDDEVQKVALFYKLDVVIDNNNTSYRLISCDTSYTTACPKISFEP	180						
Db	150	EIKKCSFRTTTSIDDEVQKVALFYKLDVVIDNNNTSYRLISCDTSYTTACPKISFEP	209						
Qy	181	IPHYCAPAGFAIILKCDKPTNGKPCGNVSTVOCTGIRVSVSTOLLNLSLBEDEVYI	240						
Db	210	IPHYCAPAGFAIILKCDKPTNGKPCGNVSTVOCTGIRVSVSTOLLNLSLBEDEVYI	269						
Qy	241	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTSKSIHIGPRAFYTTGHIIGDIRQAHNI	300						
Db	270	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTSKSIHIGPRAFYTTGHIIGDIRQAHNI	329						
Qy	301	SRACNDTLKQIVIKLAEQFENKTIIVRHSGGPEIYVMSFNGEPEFYCNSTQLPFST	360						
Db	330	SRACNDTLKQIVIKLAEQFENKTIIVRHSGGPEIYVMSFNGEPEFYCNSTQLPFST	389						
Qy	361	WNNTEGSNTTEGNTITLPCRIKQIIMWQVGRKAMTAPPRGQIRCSNTTGLLTRDG	420						
Db	390	WNNTEGSNTTEGNTITLPCRIKQIIMWQVGRKAMTAPPRGQIRCSNTTGLLTRDG	449						
Qy	421	GINENGTEIFRPGGDMRDNWRSSEFYKXKVBPLGVAPTKCRKRVQREKRAVGICAV	480						

Db 450 GINENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 509  
 Qy 481 FLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 540  
 Db 510 FLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 569  
 Qy 541 RLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 600  
 Db 570 RLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 629  
 Qy 601 TSEIYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 643  
 Db 630 TSEIYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 672

## RESULT 2

env protein - human immunodeficiency virus type 1  
 S13289  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
 Session: S13289  
 R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
 Nature 348, 63-73, 1990  
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
 A:Reference number: S13288; MUID:91043044; PMID:2172833  
 A:Accession: S13289  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-847 <OBR>  
 C:Superfamily: type E retrovirus env polyprotein

Query Match 98.1%; Score 3395; DB 2; Length 847;  
 Best Local Similarity 98.6%; Pred. No. 5.2e-237;  
 Matches 633; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 EKLMTVYVYGVPMKEATTTLFCASDAKAYDEVHNVNATHACVPTDPNQEVLNVTE 61  
 Db 31 EKLMTVYVYGVPMKEATTTLFCASDAKAYDEVHNVNATHACVPTDPNQEVLNVTE 90  
 Qy 62 HENNMKNMVEQMEQEDIIISLDOSLKPVCVLTPLCVTLNCKDVANATNTDSEGMERGE 121  
 Db 91 HENNMKNMVEQMEQEDIIISLDOSLKPVCVLTPLCVTLNCKDVANATNTDSEGMERGE 150  
 Qy 122 IKNSFNITTSIRDEVOKEVALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPI 181  
 Db 151 IKNSFNITTSIRDEVOKEVALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPI 210  
 Qy 182 PIHYCAPAGFALIKCNKDTFNGKPCPKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 241  
 Db 211 PIHYCAPAGFALIKCNKDTFNGKPCPKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 270  
 Qy 242 SDNTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGRAFYTTGIIIGDIRAHGNIS 301  
 Db 271 SDNTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGRAFYTTGIIIGDIRAHGNIS 330  
 Qy 302 RAKANDTLKOIVILRBOFEKTIIVFNHSSGSGDEIIVHNSFCGSEFFYCNSSTOLFSTW 361  
 Db 331 RAKANDTLKOIVILRBOFEKTIIVFNHSSGSGDEIIVHNSFCGSEFFYCNSSTOLFSTW 390  
 Qy 362 NNNTGSGNTEGNTITLPCRIKOIINMVEGKAMVAPIRGOIRCSNITGLLTDGG 421  
 Db 391 NNNTGSGNTEGNTITLPCRIKOIINMVEGKAMVAPIRGOIRCSNITGLLTDGG 450  
 Qy 422 INENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 481  
 Db 451 INENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 510  
 Qy 482 LGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 541  
 Db 511 LGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 570  
 Qy 542 VLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 601  
 Db 601 VLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 672

Db 571 VLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 630  
 Qy 602 SEIYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 643  
 Db 631 SEIYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 672

## RESULT 3

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)  
 T12016  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C:Accession: T12016  
 R:McCuchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.  
 AIDS Res. Hum. Retroviruses 14, 329-337, 1998  
 A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S  
 A:Reference number: Z17379; MUID:98178716; PMID:9519894  
 A:Accession: T12016  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-852 <MCC>  
 A:Cross-references: EMBL:U90934; NID:92351783; PIDN:AAC59271.1; PID:92351784  
 C:Genetic:  
 A:Gene: env  
 C:Superfamily: type E retrovirus env polyprotein

Query Match 86.9%; Score 3007.5; DB 2; Length 852;  
 Best Local Similarity 87.2%; Pred. No. 5e-209;  
 Matches 564; Conservative 31; Mismatches 47; Indels 5; Gaps 3;

Qy 2 EKLMTVYVYGVPMKEATTTLFCASDAKAYDEVHNVNATHACVPTDPNQEVLNVTE 61  
 Db 31 EKLMTVYVYGVPMKEATTTLFCASDAKAYDEVHNVNATHACVPTDPNQEVLNVTE 90  
 Qy 62 HENNMKNMVEQMEQEDIIISLDOSLKPVCVLTPLCVTLNCKDVANATNTDSEGMERGE 119  
 Db 91 HENNMKNMVEQMEQEDIIISLDOSLKPVCVLTPLCVTLNCKDVANATNTDSEGMERGE 150  
 Qy 120 GEIKNSFNITTSIRDEVOKEVALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179  
 Db 151 GEIKNSFNITTSIRDEVOKEVALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 210  
 Qy 180 PIHYCAPAGFALIKCNKDTFNGKPCPKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 239  
 Db 211 PIHYCAPAGFALIKCNKDTFNGKPCPKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 270  
 Qy 240 IRSDNTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGRAFYTTGIIIGDIRAHGN 299  
 Db 271 IRSDNTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGRAFYTTGIIIGDIRAHGN 330  
 Qy 300 ISRAKNDTLKOIVILRBOFEKTIIVFNHSSGSGDEIIVHNSFCGSEFFYCNSSTOLF 359  
 Db 331 ISRAKNDTLKOIVILRBOFEKTIIVFNHSSGSGDEIIVHNSFCGSEFFYCNSSTOLF 390  
 Qy 360 TW--NNNTGSGNTEGNTITLPCRIKOIINMVEGKAMVAPIRGOIRCSNITGLL 416  
 Db 391 TW--NNNTGSGNTEGNTITLPCRIKOIINMVEGKAMVAPIRGOIRCSNITGLL 450  
 Qy 417 TRDGINENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVG 476  
 Db 451 TRDGINENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVG 510  
 Qy 477 IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIK 536  
 Db 511 IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIK 570  
 Qy 537 QLOARVLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEERE 596  
 Db 571 QLOARVLAVERYLDGDOQLGIGWCSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEERE 630  
 Qy 597 IDNTSYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 643  
 Db 631 IDNTSYTLIESONQOEKNEOELELDKMASLWNPFDITNNLMY 677

## RESULT 4

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)  
H44001:  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Sep-1994  
C:Accession: H44001  
R:Lit., Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
J. Virol. 66, 6587-6600, 1992  
A:Title: Complete nucleotide sequence, genome organization, and biological properties of  
A:Reference number: A44001; MUID:93021387; PMID:1404605  
C:Accession: H44001  
A:Molecule type: DNA  
A:Residues: 1-843 <Lit>  
A:Cross-references: GB:M93258  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:19-35/Region: hydrophobic  
F:30-489/Product: coat protein gp120 #status predicted <GP1>  
F:490-843/Product: coat protein gp41 #status predicted <GP2>  
F:499-515/Region: hydrophobic  
F:673-689/Region: hydrophobic  
F:738-755/Domain: transmembrane #status predicted <TM>  
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 86.7%; Score 3000; DB 1; Length 843;

Best Local Similarity 87.9%; Pred. No. 1.7e-208;

Matches 566; Conservative 29; Mismatches 41; Indels 8; Gaps 6;

QY	2	EKLMTVVYGVVWMEATTTLCASDAKAYDEVNNAWTHACVPTDPNPQEVLENTE	61
DB	31	ELMTVVYGVVWMEATTTLCASDAKAYDEVNNAWTHACVPTDPNPQEVLENTE	90
QY	62	HNMMKNNVWMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGMERG	120
DB	91	HNMMKNNVWMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGMERG	150
QY	121	ELKNSFNITTSIRDEVQKEVALFYKLDVVPIDNNTSVRLSCDTSVITQACPISSEP	180
DB	151	ELKNSFNITTSIRDEVQKEVALFYKLDVVPIDNNTSVRLSCDTSVITQACPISSEP	208
QY	181	IPHYCAPAGFALIKCNDKTFNGKGPCKVSVTVOCTHGIRPVVSTOLLNGSLAEBEVI	240
DB	209	IPHYCAPAGFALIKCNDKTFNGKGPCKVSVTVOCTHGIRPVVSTOLLNGSLAEBEVI	268
QY	241	RSNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGEIIGIRQAHNI	300
DB	269	RSNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGEIIGIRQAHNI	328
QY	301	SAKXNDTLKOIVIKLREOF-ENKTIVENHSSGGDEIYTHSFNCGEFFYCNSTOLF--	359
DB	339	STQWENTLEQIAIKLEQFGNNKTIIFNPSGGDEIYTHSFNCGEFFYCNSTOLF--	386
QY	360	TNNNTGSSNTEGNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRGSNTITGLLTD	419
DB	387	TY-NOTRIKANT-GRITTLPCRIKOIIMMOEYKAMVAPPIRGQIRGSNTITGLLTD	444
QY	420	GGINENGTEIIPRGGGDMNDNRSEFYKVKVIEPLGVAPTCKKRVVOREKRAVIGA	479
DB	445	GGKDTNGTEIFRPGGDMNDNRSELYKVKVIEPLGVAPTKAKRVVOREKRAVIGA	504
QY	480	VRLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRMQLTVWGKQO	539
DB	505	LPLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRMQLTVWGKQO	564
QY	540	ARVLAVERYLDQOLLGIGWCSGKLICTAVPWNASMSKSLDRINNNMTMWEEREIDN	599

## RESULT 5

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)  
6001 YTHSYITLIESONQOEKNEBLELDDKMAISLWNPDTITNLMY 643  
DB 625 YTHSYITLIESONQOEKNEBLELDDKMAISLWNPDTITNLMY 668  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03975  
R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985  
A:Title: Nucleotide sequence of the AIDS virus, LAV.  
A:Reference number: A03975; MUID:85099333; PMID:2981635  
C:Accession: A03975  
A:Molecule type: DNA  
A:Residues: 1-861 <WAI>  
A:Cross-references: GB:K02013; NID:9326417; PID:9326424  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:517-861/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411  
F:616,621,630,642,679,755,821/Binding site: carbohydrate (asn) (covalent) #status predic

Query Match 86.4%; Score 2992.5; DB 1; Length 861;

Best Local Similarity 86.7%; Pred. No. 6.1e-208;

Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

QY	2	EKLMTVVYGVVWMEATTTLCASDAKAYDEVNNAWTHACVPTDPNPQEVLENTE	61
DB	32	EKLMTVVYGVVWMEATTTLCASDAKAYDEVNNAWTHACVPTDPNPQEVLENTE	91
QY	62	HNMMKNNVWMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGMERG	115
DB	92	HNMMKNNVWMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGMERG	151
QY	116	TMERGEIKNSFNITTSIRDEVQKEVALFYKLDVVPIDNNTSVRLSCDTSVITQACP	174
DB	152	MMERGEIKNSFNITTSIRDEVQKEVALFYKLDVVPIDNNTSVRLSCDTSVITQACP	211
QY	175	KISFEPPIHYCAPAGFALIKCNDKTFNGKGPCKVSVTVOCTHGIRPVVSTOLLNGSLA	234
DB	212	KISFEPPIHYCAPAGFALIKCNDKTFNGKGPCKVSVTVOCTHGIRPVVSTOLLNGSLA	271
QY	235	EEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGEIIGIRQAHNI	292
DB	272	EEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGEIIGIRQAHNI	330
QY	293	IRQAHNISRAKXNDTLKOIVIKLREOF-ENKTIVENHSSGGDEIYTHSFNCGEFFYCNSTOLF--	351
DB	331	IRQAHNISRAKXNDTLKOIVIKLREOF-ENKTIVENHSSGGDEIYTHSFNCGEFFYCNSTOLF--	390
QY	352	NTQTFNNTGNT--NTGSSNTEG-NTITLPCRIKOIIMMOEYKAMVAPPIRGQIRGSNTITGLLTD	407
DB	391	NTQTFNNTGNT--NTGSSNTEG-NTITLPCRIKOIIMMOEYKAMVAPPIRGQIRGSNTITGLLTD	450
QY	408	SSNITGLLTDGGINENGTEIIPRGGGDMNDNRSEFYKVKVIEPLGVAPTCKKRVVOREKRAVIGA	467
DB	451	SSNITGLLTDGGINENGTEIIPRGGGDMNDNRSELYKVKVIEPLGVAPTKAKRVVOREKRAVIGA	510
QY	468	VOREKRAVIGAVLPLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRMQLTVWGKQO	527
DB	511	VOREKRAVIGAVLPLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRMQLTVWGKQO	570



QY 533 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592  
DB 571 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 630  
QY 593 WEREIDNTSEIYTLIEESQNOEKNEQBELLELDKMAISLMMNFDTITNMLMY 643  
DB 631 WDREINNTSLIHSIEESQNOEKNEQBELLELDKMAISLMMNFDTITNMLMY 681

RESULT 8  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)

N.Alternate names: coat polyprotein  
C.Species: human immunodeficiency virus type 1, HIV-1  
A.Note: Host Homo sapiens (man)  
C.Date: 17-May-1985 #sequence\_revision 17-May-1985 #ext\_change 16-Jul-1999  
C.Accession: A03974  
R.Muehling, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313, 450-458, 1985  
A.Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus  
A.Reference number: A93355; MUID:85111157; PMID:2982104  
A.Accession: A03974  
A.Molecule type: DNA  
A.Residues: 1-856 <MUB>  
A.Cross-references: GB:K02083; NID:9555008; PIDN:AAB59873.1; PID:9328559  
C.Genetic88

A.Gene: env  
C.Superfamily: type E retrovirus env polyprotein  
C.Keywords: Aids; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F.1-30/Domain: -signal sequence #status predicted <Sig>  
F.31-51/Product: exterior membrane glycoprotein #status predicted <TM>  
F.512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F.88,136,143,156,186,197,230,234,241,282,276,289,295,301,332,339,356,386,392,397,406  
F.611,616,625,637,674,750,816/Binding site: carbohydrate (asn) (covalent) #status predic

Query Match 85.9%; Score 2974; DB 1; Length 856;  
Best Local Similarity 86.3%; Pred. No. 1.3e-206;  
Matches 562; Conservative 34; Mismatches 45; Indels 10; Gaps 6;  
QY 2 EKLWTVVYGVVWKEATTTLFCASDAKAYDEVNHNWATHACVPTDPNPOEVLNTE 61  
DB 32 EKLWTVVYGVVWKEATTTLFCASDAKAYDEVNHNWATHACVPTDPNPOEVLNTE 91  
QY 62 HNNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 119  
DB 92 HNNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 151  
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179  
DB 152 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 211  
QY 180 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVNSTOLLNGLAEEV 239  
DB 212 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVNSTOLLNGLAEEV 271  
QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 297  
DB 272 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 330  
QY 298 CNISAKNNDTLKOIVIKLREOF-ENKTIIVNHSFGDPEIIVMHSFNGEGEFYCNSTOL 356  
DB 331 CNISAKNNDTLKOIVIKLREOF-ENKTIIVNHSFGDPEIIVMHSFNGEGEFYCNSTOL 390  
QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 412  
DB 391 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 450  
QY 413 GLLTRDGGINENGIEIFRPGGDMRDNRSEFYKVKVIEPLGVAATPKARRVVOREK 472  
DB 451 GLLTRDGGINENGIEIFRPGGDMRDNRSEFYKVKVIEPLGVAATPKARRVVOREK 510  
QY 473 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAQRMQLTV 532

DB 511 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAQRMQLTV 570  
QY 533 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592  
DB 571 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 630  
QY 593 WEREIDNTSEIYTLIEESQNOEKNEQBELLELDKMAISLMMNFDTITNMLMY 643  
DB 631 WDREINNTSLIHSIEESQNOEKNEQBELLELDKMAISLMMNFDTITNMLMY 681

RESULT 9

env protein - human immunodeficiency virus type 1  
S13288  
C.Species: human immunodeficiency virus type 1, HIV-1  
C.Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #ext\_change 31-Oct-1997  
C.Accession: S13288  
R.O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
Nature 348, 69-73, 1990  
A.Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
A.Reference number: S13288; MUID:91043044; PMID:2172833  
A.Accession: S13288  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-854 <OBR>  
C.Superfamily: type E retrovirus env polyprotein

Query Match 85.8%; Score 2972; DB 2; Length 854;  
Best Local Similarity 86.6%; Pred. No. 1.8e-206;  
Matches 564; Conservative 30; Mismatches 45; Indels 12; Gaps 7;  
QY 2 EKLWTVVYGVVWKEATTTLFCASDAKAYDEVNHNWATHACVPTDPNPOEVLNTE 61  
DB 32 EKLWTVVYGVVWKEATTTLFCASDAKAYDEVNHNWATHACVPTDPNPOEVLNTE 91  
QY 62 HNNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 119  
DB 92 HNNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 151  
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179  
DB 152 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 209  
QY 180 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVNSTOLLNGLAEEV 239  
DB 210 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVNSTOLLNGLAEEV 269  
QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 297  
DB 270 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 328  
QY 298 CNISAKNNDTLKOIVIKLREOF-ENKTIIVNHSFGDPEIIVMHSFNGEGEFYCNSTOL 356  
DB 329 CNISAKNNDTLKOIVIKLREOF-ENKTIIVNHSFGDPEIIVMHSFNGEGEFYCNSTOL 388  
QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 412  
DB 389 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 448  
QY 413 GLLTRDGGINENGIEIFRPGGDMRDNRSEFYKVKVIEPLGVAATPKARRVVOREK 472  
DB 449 GLLTRDGGINENGIEIFRPGGDMRDNRSEFYKVKVIEPLGVAATPKARRVVOREK 508  
QY 473 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAQRMQLTV 532  
DB 509 RAVGIGAVFLGFLGAAGSTMGCTSWTLTVQARLLLSGIVQOQNLLRAIEAQRMQLTV 568  
QY 533 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592  
DB 569 WGIKOLARVLAVERYLGDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 628  
QY 593 WEREIDNTSEIYTLIEESQNOEKNEQBELLELDKMAISLMMNFDTITNMLMY 643

Db 629 WDREINNTSLIHSLSIESQNOEKNEBQELLELDPMKASLMMFNTLWMTY 679

RESULT 10  
VCLJAJ2  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03976  
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-SH  
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).  
A:Reference number: A04003; MUID:85090453; PMID:2578227  
A:Accession type: DNA  
A:Molecule type: A03976  
A:Residues: 1-855 <SPAN>  
A:References: GB:K02007; NID:G328658; PIDN:AAB59882.1; PID:G328666  
A:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>  
F:87,129,140,158,184,190,200,244,265,282,298,304,334,351,358,364,388,394,400,408,445,458  
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status Predicted

Query Match 85.8%; Score 2972; DB 1; Length 855;  
Best Local Similarity 86.2%; Pred. No. 1.8e-206;  
Matches 563; Conservative 32; Mismatches 44; Indels 14; Gaps 8;

Qy 2 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61  
Db 31 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90  
Qy 62 HFNMKNMNMVQOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDVANTNTNDSGTM-E-RG 120  
Db 91 HFNMKNMNMVQOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDVANTNTNDSGTM-E-RG 150  
Qy 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVPIDN-----NNTSYRLISCDTSVITQACPK 175  
Db 151 EIKNCSFNITTSIRDEVOKEYALFYKLDVPIDN-----NNTSYRLISCDTSVITQACPK 210  
Qy 176 ISFEPDPIHYCAPAGFALIKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAE 235  
Db 211 ISFEPDPIHYCAPAGFALIKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAE 270  
Qy 236 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 295  
Db 271 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 330  
Qy 296 AHCNISRAKMDTLKQIYIKLEQFENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 354  
Db 331 AHCNISRAKMDTLKQIYIKLEQFENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 390  
Qy 355 QLFNSTWNNN---TESSNNT-EGNTTTLPCRIOIINMGOEVKANYAPPIRQIICSS 412  
Db 391 QLFNSTWNNN---TESSNNT-EGNTTTLPCRIOIINMGOEVKANYAPPIRQIICSS 447  
Qy 413 GLLLTRDGGIN-ENKTEIFRPGGDMRDNRSEFYKVKVIEPLGVAFTCKRRVORE 471  
Db 448 GLLLTRDGGIN-ENKTEIFRPGGDMRDNRSEFYKVKVIEPLGVAFTCKRRVORE 507  
Qy 472 KRAVGI-GAVFLGFLGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRMQL 530  
Db 508 KRAVGI-GAVFLGFLGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRMQL 567  
Qy 531 TWVGIKQIQARVLAVERLYGQOQLLGIVGSGGKLICTTAVPWNASWSKSLDRINWNTW 590  
Db 568 TWVGIKQIQARVLAVERLYGQOQLLGIVGSGGKLICTTAVPWNASWSKSLDRINWNTW 627

Qy 591 MEMEREIDNTSYTYLLESQNOEKNEBQELLELDPMKASLMMFNTLWMTY 643  
Db 628 MEMEREIDNTSYTYLLESQNOEKNEBQELLELDPMKASLMMFNTLWMTY 680

RESULT 11  
VCLJMN  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)

N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C:Accession: A28922  
R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Stat  
A:Title: Envelope sequences of two new United States HIV-1 isolates.  
A:Reference number: A28922; MUID:88219542; PMID:3369091  
A:Accession type: DNA  
A:Molecule type: A28922  
A:Residues: 1-859 <CUR>  
A:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-859/Product: env polyprotein #status predicted <EP>  
F:87,129,135,140,141,146,161,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401

Query Match 85.7%; Score 2967.5; DB 1; Length 859;  
Best Local Similarity 85.7%; Pred. No. 3.9e-206;  
Matches 562; Conservative 32; Mismatches 47; Indels 15; Gaps 6;

Qy 2 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61  
Db 31 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90  
Qy 62 HFNMKNMNMVQOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTN-----DS 113  
Db 91 HFNMKNMNMVQOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTN-----DS 150  
Qy 114 EGTMERSEIKNCSFNITTSIRDEVOKEYALFYKLDVPIDNNTSYRLISCDTSVITQAC 173  
Db 151 EGTMERSEIKNCSFNITTSIRDEVOKEYALFYKLDVPIDNNTSYRLISCDTSVITQAC 210  
Qy 174 PKISFEPDPIHYCAPAGFALIKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSL 233  
Db 211 PKISFEPDPIHYCAPAGFALIKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSL 270  
Qy 234 AEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 293  
Db 271 AEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 330  
Qy 294 ROAHNISRAKMDTLKQIYIKLEQFENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 353  
Db 331 ROAHNISRAKMDTLKQIYIKLEQFENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 390  
Qy 354 TQLNSTWNNN---TESSNNT-EGNTTTLPCRIOIINMGOEVKANYAPPIRQIICSS 409  
Db 391 TQLNSTWNNN---TESSNNT-EGNTTTLPCRIOIINMGOEVKANYAPPIRQIICSS 450  
Qy 410 NITGLLTRDGG--INENGEIFRPGGDMRDNRSEFYKVKVIEPLGVAFTCKRRV 467  
Db 451 NITGLLTRDGG--INENGEIFRPGGDMRDNRSEFYKVKVIEPLGVAFTCKRRV 510  
Qy 468 VOREKRAVGI-GAVFLGFLGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRM 527  
Db 511 VOREKRAVGI-GAVFLGFLGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRM 569  
Qy 528 LQLTWVGIKQIQARVLAVERLYGQOQLLGIVGSGGKLICTTAVPWNASWSKSLDRINWNTW 587  
Db 570 LQLTWVGIKQIQARVLAVERLYGQOQLLGIVGSGGKLICTTAVPWNASWSKSLDRINWNTW 629



QY 588 MTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITWMLWY 643  
 Db 630 MTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITWMLWY 685

## RESULT 12

VCLJ3W

env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)  
 N:Alternate names: coat protein  
 C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997

C:Accession: B28922

R:Guigo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sci

A:Title: Envelope sequences of two new United States HIV-1 isolates.

A:Reference number: A28922; MUID:88219542; PMID:3369091

A:Accession: B28922

A:Molecule type: DNA

A:Residues: 1-861 <GUR>

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane prote

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-861/Product: env polyprotein #status predicted <EP>

F:87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 85.7%; Score 2967; DB 1; Length 861;

Best Local Similarity 85.6%; Pred. No. 4.2e-206;

Matches 565; Conservative 34; Mismatches 39; Indels 22; Gaps 9;

QY 2 EKLMTVTVYGVVPMKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENT 61  
 Db 31 EQLMTVTVYGVVPMKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENT 90

QY 62 HENMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSRG 115  
 Db 91 HENMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSRG 150

QY 116 TMEBEIKKCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 175  
 Db 151 KMEGEMNTNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 208

QY 176 ISFEPIPIHYCAPAGFALIKCNDKTFNKGPCCKANSTVQCTHGIRPVYSTOLLNGSLAE 235  
 Db 209 VSFEPIPIHYCAPAGFALIKCNDKTFNKGPCCKANSTVQCTHGIRPVYSTOLLNGSLAE 267

QY 236 EEVVRISDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIGDI 293  
 Db 268 EEVVRISDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFTTGEIIGDI 327

QY 294 QAHNCNISRAKNNNTLKQIVIKLRQEFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 353  
 Db 328 QAHNCNISRAKNNNTLKQIVIKLRQEFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 387

QY 354 TOLFSTNNNTNTEG---NTEGNT--TITLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 409  
 Db 388 TOLFSTNNNTNTEG---NTEGNT--TITLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 446

QY 410 NITGLLITRDGINENG---TEIFRPGGDMRDMMSEFFYKVKVIEPLGVAFTPKC 464  
 Db 447 NITGLLITRDGINENG---TEIFRPGGDMRDMMSEFFYKVKVIEPLGVAFTPKC 506

QY 465 RRVVREKRAVNG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEA 523  
 Db 507 RRVVREKRAVNG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEA 566

QY 524 QORMLQITWGIKQIOLARVLAVERVYLDQDQGLIGWCSGKLCCTAVPNAASNSKSIDR 583  
 Db 567 QORMLQITWGIKQIOLARVLAVERVYLDQDQGLIGWCSGKLCCTAVPNAASNSKSIDR 626

QY 584 IWNMTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITWMLWY 643

Db 627 IWNMTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITWMLWY 686

## RESULT 13

VCLJ3W

env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)

N:Contains: coat protein gp120; coat protein gp41

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 07-Nov-1997

C:Accession: A24774

R:Starcich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E

Cell 45, 637-648, 1986

A:Title: Identification and characterization of conserved and variable regions in the en

A:Reference number: A24774; MUID:86218077; PMID:2423250

A:Accession: A24774

A:Molecule type: DNA

A:Residues: 1-856 <STA>

A:Cross-references: GB:K03455; GB:M38432; NID:gi906382

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-501/Product: coat protein gp120 #status predicted <GP1>

F:502-847/Product: coat protein gp41 #status predicted <GP2>

F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match 85.5%; Score 2958.5; DB 1; Length 856;

Best Local Similarity 86.1%; Pred. No. 1.7e-205;

Matches 564; Conservative 30; Mismatches 46; Indels 15; Gaps 7;

QY 1 VEKLMVTYVGVVPMKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENT 60  
 Db 30 VEKLMVTYVGVVPMKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENT 89

QY 61 EHFNMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATVNTNDSRGTWERG 120  
 Db 90 EHFNMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATVNTNDSRGTWERG 147

QY 121 EIKNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 175  
 Db 148 EIKNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 207

QY 176 ISFEPIPIHYCAPAGFALIKCNDKTFNKGPCCKANSTVQCTHGIRPVYSTOLLNGSLAE 235  
 Db 208 VSFEPIPIHYCAPAGFALIKCNDKTFNKGPCCKANSTVQCTHGIRPVYSTOLLNGSLAE 267

QY 236 EEVVRISDNFTNNAKTIIVQLKESVEINCTRPNNNT-RKSIHIGPRAFTTGEIIGDI 294  
 Db 268 EEVVRISDNFTNNAKTIIVQLKESVEINCTRPNNNT-RKSIHIGPRAFTTGEIIGDI 326

QY 295 QAHNCNISRAKNNNTLKQIVIKLRQEFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 354  
 Db 327 QAHNCNISRAKNNNTLKQIVIKLRQEFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 386

QY 355 OLFNSTNN---NTEGNTNTE--GNTITLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 409  
 Db 387 OLFNSTNN---NTEGNTNTE--GNTITLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 446

QY 410 NITGLLITRDGINENGTEIFRPGGDMRDMMSEFFYKVKVIEPLGVAFTPKCGRVVO 469  
 Db 447 NITGLLITRDGINENGTEIFRPGGDMRDMMSEFFYKVKVIEPLGVAFTPKCGRVVO 506

QY 470 REKRAVG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQOQML 528  
 Db 507 REKRAVG--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQOQML 566

QY 529 QITWGIKQIOLARVLAVERVYLDQDQGLIGWCSGKLCCTAVPNAASNSKSIDR 588  
 Db 567 QITWGIKQIOLARVLAVERVYLDQDQGLIGWCSGKLCCTAVPNAASNSKSIDR 626

QY 589 IWNMTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITWMLWY 643

Db 627 TWMEWERIDNVTSLIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 681

## RESULT 14

S33985

env polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33985

R:Carlini, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-851 &lt;CNR&gt;

A:Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA77628.1; PID:G60199

C:Superfamily: type E retrovirus env polyprotein

Query Match 85.3%; Score 2953.5; DB 2; Length 851;

Best Local Similarity 86.3%; Pred. No. 3.9e-205;

Matches 559; Conservative 34; Mismatches 46; Indels 9; Gaps 6;

Qy 2 EKLWTVYGVVWPKWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 61

Db 32 EKLWTVYGVVWPKWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 91

Qy 62 HFNWKNMVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTDSEGTMER 119

Db 92 NFNWKNMVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTDSEGTMER 151

Qy 120 GEINCSFNITTSIRDEVOKEVALFYKLDVPIIDNNVTYLLSCDSVITQACPKE 179

Db 152 GEINCSFNITTSIRDEVOKEVALFYKLDVPIIDNNVTYLLSCDSVITQACPKE 211

Qy 180 PIPHYCAPAGFALCKNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGSLAESEV 239

Db 212 PIPHYCAPAGFALCKNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGSLAESEV 271

Qy 240 IRSDFNNTAKTIIVOLKESVEINCTRPNNTRKSIH-GRGAFYTTGSIIGDIRQAH 297

Db 272 IRSDFNNTAKTIIVOLKESVEINCTRPNNTRKSIH-GRGAFYTTGSIIGDIRQAH 330

Qy 298 CNISRAKNDTLKOIVIKLREOF-BNKTIIVNHSGGDPEIVMHSFNGEGFFYCNSTOL 356

Db 331 CNISRAKNDTLKOIVIKLREOF-BNKTIIVNHSGGDPEIVMHSFNGEGFFYCNSTOL 390

Qy 357 FNSTNNNTGSSNNTG-NTITLPCRIKOIIMMOEVGKAMVAPPIRGQIRCSSNITGL 415

Db 391 FNSTNNNTGSSNNTG-NTITLPCRIKOIIMMOEVGKAMVAPPIRGQIRCSSNITGL 448

Qy 416 LTRDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 475

Db 449 LTRDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 508

Qy 476 GIGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 535

Db 509 GIGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 568

Qy 536 KQOLARVLAVERYLQDQOLGIMGSGKLTCTTAVPNAASNSKSLDRINMMTMWEMER 595

Db 569 KQOLARVLAVERYLQDQOLGIMGSGKLTCTTAVPNAASNSKSLDRINMMTMWEMER 628

Qy 596 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 643

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 629 EIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

N:Contains: coat protein gp120; coat protein gp32

C:Species: human immunodeficiency virus type 1, HIV-1

A:Name: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 01-Mar-1996

C:Accession: B42995

R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Uehijima, H.; Kitamura, T.

Virology 189, 534-546, 1992

A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated

A:Reference number: A42995; MUID:92351552; PMID:1322587

A:Accession: B42995

A:Molecule type: mRNA

A:Residues: 1-729 &lt;SHI&gt;

A:Cross-references: GB:S41266; GB:D01206

C:Genetics:

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-689/Domain: extracellular #status predicted &lt;EXT&gt;

F:1-33/Domain: signal sequence #status predicted &lt;SIG&gt;

F:17-33/Region: hydrophobic #status predicted

F:34-517/Product: coat protein gp120 #status predicted &lt;CP1&gt;

F:514-517/Region: cleavage processing #status predicted &lt;CP2&gt;

F:518-729/Product: coat protein gp32 #status predicted &lt;CP2&gt;

F:518-534/Region: hydrophobic #status predicted

F:690-711/Domain: transmembrane #status predicted &lt;TM1&gt;

F:712-729/Domain: intracellular #status predicted &lt;INT&gt;

F:93,141,145,146,163,191,192,227,241,248,269,283,296,308,338,345,361,367,397,403,408,414,

Query Match 83.6%; Score 2895.5; DB 1; Length 729;

Best Local Similarity 82.8%; Pred. No. 4.2e-201;

Matches 539; Conservative 40; Mismatches 63; Indels 9; Gaps 2;

Qy 2 EKLWTVYGVVWPKWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 61

Db 37 EKLWTVYGVVWPKWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 96

Qy 62 HFNWKNMVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTDSEGTMER 121

Db 97 HFNWKNMVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTDSEGTMER 156

Qy 122 IKNCSEFNITTSIRDEVOKEVALFYKLDVPIIDNN-NTSYLLSCDSVITQACPKE 177

Db 157 IKNCSEFNITTSIRDEVOKEVALFYKLDVPIIDNNNTSYLLSCDSVITQACPKE 216

Qy 178 PIPHYCAPAGFALCKNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGSLAESEV 237

Db 217 PIPHYCAPAGFALCKNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGSLAESEV 276

Qy 238 VIRSDFNNTAKTIIVOLKESVEINCTRPNNTRKSIH-GRGAFYTTGSIIGDIRQAH 297

Db 277 VIRSDFNNTAKTIIVOLKESVEINCTRPNNTRKSIH-GRGAFYTTGSIIGDIRQAH 336

Qy 298 CNISRAKNDTLKOIVIKLREOF-BNKTIIVNHSGGDPEIVMHSFNGEGFFYCNSTOL 357

Db 337 CNISRAKNDTLKOIVIKLREOF-BNKTIIVNHSGGDPEIVMHSFNGEGFFYCNSTOL 396

Qy 358 NSTNNNTGSSNNTG-NTITLPCRIKOIIMMOEVGKAMVAPPIRGQIRCSSNITGL 417

Db 397 NSTNNNTGSSNNTG-NTITLPCRIKOIIMMOEVGKAMVAPPIRGQIRCSSNITGL 456

Qy 418 RDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 472

Db 457 RDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 516

Qy 473 RAVGIGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 532

Db 517 RAVGIGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 576

Qy 533 MGKQOLARVLAVERYLQDQOLGIMGSGKLTCTTAVPNAASNSKSLDRINMMTMWEMER 592

Db 577 MGKQOLARVLAVERYLQDQOLGIMGSGKLTCTTAVPNAASNSKSLDRINMMTMWEMER 636

Qy 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 643

Db 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Qy 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 593 WEREIDNVTSEIYTLIESQNOQEKNEQELLELDKMASLNMWFTITWLMWY 676

Db 637 WEREINNYTNLYNLIBESQOQEKNEODLALDKWDSLNNWFSITKXLMY 687

Search completed: December 12, 2003, 12:33:31  
Job time: 16.9697 secs

**THIS PAGE BLANK (US:70)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:25:38 ; Search time 10.7921 Seconds  
(without alignments)  
2801.881 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYYGVVPMKEATL.....ELDKWASIMNFDITNLMY 643

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

1 number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3139.5	90.7	848	1 ENV_HV1JR	P20871 human immun
2	3041	87.8	847	1 ENV_HV1S1	P19550 human immun
3	3027.5	87.4	856	1 ENV_HV1SC	P05878 human immun
4	3000	86.7	843	1 ENV_HV1Y2	P35961 human immun
5	2997	86.6	867	1 ENV_HV1J3	P12489 human immun
6	2993.5	86.4	861	1 ENV_HV1BR	P03377 human immun
7	2992	86.4	852	1 ENV_HV1BN	P12488 human immun
8	2978	86.0	856	1 ENV_HV1BN	P03375 human immun
9	2974	85.9	855	1 ENV_HV1PV	P03376 human immun
10	2972	85.8	855	1 ENV_HV1A2	P03378 human immun
11	2964	85.6	856	1 ENV_HV1H2	P04578 human immun
12	2958	85.5	856	1 ENV_HV1W1	P31872 human immun
13	2954.5	85.3	851	1 ENV_HV1H3	P04624 human immun
14	2954.5	85.2	851	1 ENV_HV1S3	P19549 human immun
15	2948.5	85.2	851	1 ENV_HV1B8	P04582 human immun
16	2946	85.1	856	1 ENV_HV1LW	O70626 human immun
17	2946	85.1	856	1 ENV_HV1LW	O70626 human immun
18	2917	84.3	855	1 ENV_HV1OY	P19551 human immun
19	2917	84.3	855	1 ENV_HV1OY	P20888 human immun
20	2910	84.1	847	1 ENV_HV1R2	P04579 human immun
21	2895.5	83.6	861	1 ENV_HV1K8	P05880 human immun
22	2895.5	83.6	861	1 ENV_HV1K8	P31819 human immun
23	2895.5	83.6	861	1 ENV_HV1C4	P05879 human immun
24	2706	78.2	855	1 ENV_HV1Z6	P04580 human immun
25	2695	77.8	863	1 ENV_HV1Z8	P05882 human immun
26	2686	77.6	853	1 ENV_HV1Z2	P12487 human immun
27	2657.5	76.8	859	1 ENV_HV1MA	P04583 human immun
28	2638.5	76.2	856	1 ENV_HV1Z8	P05881 human immun
29	2614.5	75.5	846	1 ENV_HV1ND	P18799 human immun
30	2614	75.5	853	1 ENV_HV1EL	P04581 human immun
31	2244	64.8	854	1 ENV_SIVC2	P17281 chimpanzee
32	1721.5	49.7	421	1 ENV_HV1N5	P12490 human immun
33	1613	46.6	460	1 ENV_HV1Z3	P12491 human immun

## ALIGNMENTS

RESULT 1	ID	ENV_HV1JR	STANDARD	PRT	848 AA.	
AC	P20871					Q02837 simian immu
DT	01-FEB-1991 (Rel. 17, Last sequence update)					P20872 human immun
DT	15-SEP-2003 (Rel. 42, Last annotation update)					P18040 human immun
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].					P27757 simian immu
GN	ENV					P32536 human immun
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).					P04577 human immun
OC	Viruses; Retroviridae; Retroviridae; Lentivirus.					P27977 simian immu
OX	NCBI_TaxID=11688;					P05886 simian immu
RN	SEQUENCE FROM N.A.					P24105 human immun
RA	Koyanagi S., Chen I.S.Y.;					P05883 human immun
RL	Submitted (DEC-1988) to the HIV data bank.					P17755 human immun
CC	-----					P12449 human immun
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
CC	EMBL: M38429; AAB03749.1; ..					
DR	PDB: 1CE4; 1B-MAR-99.					
DR	HIV; M38429; ENV5URCSF.					
DR	InterPro: IPR000328; Env_GP41.					
DR	InterPro: IPR000777; GP120.					
DR	Pfam: PF00516; GP120; 1.					
DR	Pfam: PF00517; GP41; 1.					
KW	AIDS; Coat protein; Polypeptide; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.					
KW	SIGNAL	1	32			
FT	CHAIN	33	503			EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	848			TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73			BY SIMILARITY.
FT	DISULFID	118	203			BY SIMILARITY.
FT	DISULFID	125	194			BY SIMILARITY.
FT	DISULFID	130	154			BY SIMILARITY.
FT	DISULFID	216	245			BY SIMILARITY.
FT	DISULFID	226	237			BY SIMILARITY.
FT	DISULFID	294	328			BY SIMILARITY.
FT	DISULFID	374	437			BY SIMILARITY.
FT	DISULFID	381	410			BY SIMILARITY.
FT	CARBOHYD	87	87			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	134	134			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	137	137			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	153	153			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157	157			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	185	185			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	228	228			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239			N-LINKED (GLCNAC. . .) (POTENTIAL).

AC	PL19550: 01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DT	Envelope polypeptide GP160 precursor (contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
DE	ENV.
GN	Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OS	Vitruves; Retroid vitruves; Retroviridae; Lentivirus.
OC	NCBI_TaxID=11691;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=90347835; PubMed=2384920;
RX	Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.
RA	"Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RT	J. Virol. 64:4390-4390(1990).
RL	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
CC	EMBL; M65024; AAA45072.1; -
DR	PDB; LOBE; 15-MAY-97.
DR	HIV; M38428; ENV5SF162.
DR	InterPro; IPR000326; Env GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.
KM	
FT	SIGNAL 1 29
FT	CHAIN 30 502
FT	CHAIN 503 847
FT	DISULFID 53 73
FT	DISULFID 118 203
FT	DISULFID 125 194
FT	DISULFID 130 155
FT	DISULFID 216 245
FT	DISULFID 226 237
FT	DISULFID 294 325
FT	DISULFID 374 435
FT	DISULFID 381 408
FT	CARBOHYD 87 408
FT	CARBOHYD 135 135
FT	CARBOHYD 154 154
FT	CARBOHYD 186 186
FT	CARBOHYD 195 195
FT	CARBOHYD 232 232
FT	CARBOHYD 239 239
FT	CARBOHYD 260 260
FT	CARBOHYD 274 274
FT	CARBOHYD 293 293
FT	CARBOHYD 299 299
FT	CARBOHYD 329 329
FT	CARBOHYD 336 336
FT	CARBOHYD 352 352
FT	CARBOHYD 382 382
FT	CARBOHYD 388 388
FT	CARBOHYD 392 392
FT	CARBOHYD 398 398
FT	CARBOHYD 401 401
FT	CARBOHYD 438 438
FT	CARBOHYD 454 454
FT	CARBOHYD 602 602
FT	CARBOHYD 607 607
FT	CARBOHYD 616 616
FT	CARBOHYD 628 628
FT	SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

Query Match 87.8%; Score 3041; DB 1; Length 847;  
 Best Local Similarity 88.7%; Pred. No. 5.9e-228;  
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

QY 1 VEKMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLENVT 60  
 DB VEKMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLENVT 89  
 QY 61 EHFNNKNNWVQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGSTER 119  
 DB EHFNNKNNWVQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGSTER 149  
 QY 90 EHFNNKNNWVQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGSTER 149  
 DB 90 EHFNNKNNWVQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGSTER 149  
 QY 120 GEIKNCSFNTTISRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 179  
 DB 120 GEIKNCSFNTTISRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 209  
 QY 150 GEIKNCSFNTTISRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 209  
 DB 150 GEIKNCSFNTTISRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSVITQACPKISFE 209  
 QY 180 PIPHYCAPAGFAILLKNDKTFNGKGPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 239  
 DB 210 PIPHYCAPAGFAILLKNDKTFNGKGPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 269  
 QY 240 IRSDFPTNNAKTIYQLESVEINCTRPNNTRKSIHIGRGAFTTGTGIIIGDIOAHON 299  
 DB 270 IRSDFPTNNAKTIYQLESVEINCTRPNNTRKSIHIGRGAFTTGTGIIIGDIOAHON 329  
 QY 300 ISRAKNDTKQIVIKLREOPENKTIYFNHSSGDPPEIYMHSGEPEFYCNSTOLFNS 359  
 DB 330 ISRAKNDTKQIVIKLREOPENKTIYFNHSSGDPPEIYMHSGEPEFYCNSTOLFNS 389  
 QY 350 TNNNTTSGSNTTGTITLPCRIKQIINMOEYKAMVAPPIRGOIRCSNITGLLTRD 419  
 DB 390 TNNNTTSGSNTTGTITLPCRIKQIINMOEYKAMVAPPIRGOIRCSNITGLLTRD 447  
 QY 420 GGIN-ENGTEIRRCGGDMRWSEFYKYKTKIPLGVATTKCKRRVVOEKAVGIG 478  
 DB 448 GKEISNTTELFRPGGDMRWSEFYKYKTKIPLGVATTKCKRRVVOEKAVGIG 507  
 QY 479 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQOQRLQTLWIGIKOL 538  
 DB 508 AMFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQOQRLQTLWIGIKOL 567  
 QY 539 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 598  
 DB 568 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 627  
 QY 599 NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 643  
 DB 628 NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 672

RESULT 3  
 ENV\_HVISC STANDARD; PRT; 856 AA.  
 AC P05878;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN GN  
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NC NCBI\_Taxid:11702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80219542; PubMed=3369091;  
 RA Guro C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,  
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;  
 RT "Envelope sequences of two new United States HIV-1 isolates."  
 RL Virology 144:531-536(1988).  
 CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN  
 CC 1984 IN SOUTHERN CALIFORNIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: M17450; -; NOT\_ANNOTATED\_CDS.  
 DR HIV: M17450; ENVSSC.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR00777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 510  
 FT CHAIN 511 856  
 FT SITE 760 760  
 FT DISULFID 53 73  
 FT DISULFID 118 206  
 FT DISULFID 125 197  
 FT DISULFID 130 160  
 FT DISULFID 219 247  
 FT DISULFID 228 239  
 FT DISULFID 296 330  
 FT DISULFID 376 439  
 FT DISULFID 383 412  
 FT CARBOHYD 87 87  
 FT CARBOHYD 129 129  
 FT CARBOHYD 135 135  
 FT CARBOHYD 140 140  
 FT CARBOHYD 143 143  
 FT CARBOHYD 159 159  
 FT CARBOHYD 163 163  
 FT CARBOHYD 183 183  
 FT CARBOHYD 198 198  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 302 302  
 FT CARBOHYD 331 331  
 FT CARBOHYD 338 338  
 FT CARBOHYD 354 354  
 FT CARBOHYD 360 360  
 FT CARBOHYD 384 384  
 FT CARBOHYD 394 394  
 FT CARBOHYD 400 400  
 FT CARBOHYD 405 405  
 FT CARBOHYD 442 442  
 FT CARBOHYD 457 457  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 625 625  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CMC64;

Query Match 87.4%; Score 3027.5; DB 1; Length 856;  
 Best Local Similarity 87.2%; Pred. No. 6.7e-227;  
 Matches 571; Conservative 33; Mismatches 34; Indels 17; Gaps 7;

QY 2 EKLMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLENVT 61  
 DB 31 EKLMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLENVT 90

```

Db      91 NFNWKNMVEQWHEDIISLMDOSLKPCKVLTPLCVTLNCTNLNRDSTNATNTSSNRG 150
Qy      116 TMRGEIKNGSFNTTTSIRDEVOKEVALFYKLDVDPIDNNNTSYRLISCDPSVITQACPK 175
Db      151 KMEGEMTNGSFNTTTSIRSKYKVALFYKLDVDPID--NTSTLLNCTNLSVITQACPK 208
Qy      176 ISPEPIPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAE 235
Db      209 VSFEPPIPIHYCA-RMPAILNCKNKFNKGTGCTNVSTVQCTHGIRPVVSTOLLNGSLAE 267
Qy      236 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGGRAPFTTGEIIGDIRQ 295
Db      268 EEVVLRSENFNTNNAKTIIVOLKEAVEINCTRPNNNTKRSIHIGGRAPFTTGEIIGDIRQ 327
Qy      296 AHGWSAKMNDTLKQVYKLRQEPNKTIVFNHSSGSDPEIWNHSPFCGSEFPYCNSTQ 355
Db      328 AHGWSAKMNDTLKQVYKLRQEPNKTIVFNHSSGSDPEIWNHSPFCGSEFPYCNSTQ 387
Qy      356 LFNSTNNNTGSGNNTGEGN-TITLPCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGL 414
Db      388 LFSSTW-NGTEGSGNNTGNDTITLPCRIKELIINMWOEYKAMVAPPIRGQIRCSNITGL 446
Qy      415 LITPDGGINENG-----TEIFRPGGDMRDNRSEFPYKRYKVEIPLGVAPTKCKRRVYQ 469
Db      447 LITPDGSGSKSKSKNENTEIFRPGGDMRDNRSELYKYKVKIPLGVAPTKCKRRVYQ 506
Qy      470 REKRAVG-IGAVPLGFGAAGSTMGASMTLTVQARLLSGIVQOQNNLLAIEAQORML 528
Db      507 REKRAVGITGMFGLGFGAAGSTMGATSMITLVQARLLSGIVQOQNNLLAIEAQORML 566
Qy      529 QLTWAGIKQIQLARVLAVERVYLGDOQLGIGWCSGRLICCTAVPWNASNSKSLDRIMNM 588
Db      567 QLTWAGIKQIQLARVLAVERVYLRDQQLGIGWCSGRLICTTTPVNMNTSMNSKSLDKIMNM 626
Qy      589 TMMWERIDNVTSEITVTLIESQNOEKNOELLDKASLMMWPTITWLMY 643
Db      627 TMMWERIDNVTSLITVTLIESQNOEKNOELLDKASLMMWPTITWLMY 691

RESULT 4
ENV_HV1Y2 STANDARD; PRT; 843 AA.
ID ENV_HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
ENV.
CC Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
CC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CC NCBI_TaxID=36377;
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M93258; -, NOT_ANNOTATED_CDS.
CC PIR, H44001; H44001.
CC PDB, 1G9N; 27-DEC-00.

```

```

DR      InterPro; IPR000328; Env_GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.
FT      SIGNAL 1 29
FT      CHAIN 30 489
FT      CHAIN 490 843
FT      TRANSMEM 738 755
FT      DISULFID 53 73
FT      DISULFID 118 201
FT      DISULFID 125 192
FT      DISULFID 130 155
FT      DISULFID 214 243
FT      DISULFID 224 235
FT      DISULFID 292 326
FT      DISULFID 373 432
FT      DISULFID 380 405
FT      CARBOHYD 87 87
FT      CARBOHYD 129 129
FT      CARBOHYD 135 135
FT      CARBOHYD 138 138
FT      CARBOHYD 154 154
FT      CARBOHYD 158 158
FT      CARBOHYD 184 184
FT      CARBOHYD 193 193
FT      CARBOHYD 230 230
FT      CARBOHYD 237 237
FT      CARBOHYD 258 258
FT      CARBOHYD 272 272
FT      CARBOHYD 285 285
FT      CARBOHYD 291 291
FT      CARBOHYD 297 297
FT      CARBOHYD 327 327
FT      CARBOHYD 351 351
FT      CARBOHYD 381 381
FT      CARBOHYD 389 389
FT      CARBOHYD 395 395
FT      CARBOHYD 400 400
FT      CARBOHYD 435 435
FT      CARBOHYD 450 450
FT      CARBOHYD 598 598
FT      CARBOHYD 603 603
FT      CARBOHYD 612 612
FT      CARBOHYD 624 624
FT      CARBOHYD 803 803
SQ      SEQUENCE 843 AA; 95648 MW; C69DFD971C918971 CRC64;

Query Match      86.7%; Score 3000; DB 1; Length 843;
Best Local Similarity 87.9%; Pred. No. 8.9e-225;
Matches 566; Conservative 29; Mismatches 41; Indels 8; Gaps 6;

Qy      2 EKLMTVYVYGVYVWKEATTTLLFCASDAKAYDTEVHNVMAHACVPTDPNPOEVVLENT 61
Db      31 EQLMTVYVYGVYVWKEATTTLLFCASDAKAYDTEVHNVMAHACVPTDPNPOEVVLENT 90
Qy      62 HENWKNMVEQWHEDIISLMDOSLKPCKVLTPLCVTLNCTNLNRDSTNATNTSSNRG 120
Db      91 HENWKNMVEQWHEDIISLMDOSLKPCKVLTPLCVTLNCTNLNRDSTNATNTSSNRG 150
Qy      121 EIKKCSFNITTSIRDEVOKEVALFYKLDVDPIDNNNTSYRLISCDPSVITQACPKISFEP 180
Db      151 EIKKCSFNITTSIRDKYKVALFYKLDVDPID--NASTYLLSNTSVITQACPKVSFEP 208
Qy      181 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVVI 240
Db      209 IPIHYCAPAGFAILKCNKDKFNGKGPCTNVSTVQCTHGIRPVVSTOLLNGSLAESEVVI 268
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGGRAPFTTGEIIGDIRQAHGNI 300
Db      269 RSENFNTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGGRAPFTTGEIIGDIRQAHGNI 328

```





Query Match	Best Local Similarity	86.4%;	Score 2992.5;	DB 1;	Length 861;
Matches 569;	Conservative	32;	Mismatches 40;	Indels 15;	Gaps 8
FT CARBOHYD	239	239	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	246	246	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	267	267	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	281	281	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	294	294	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	300	300	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	306	306	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	337	337	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	344	344	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	361	361	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	391	391	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	397	397	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	402	402	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	411	411	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	453	453	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	468	468	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	616	616	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	621	621	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	630	630	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	642	642	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	679	679	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	755	755	N-LINKED (GLCNAC .)	(POTENTIAL)	-
FT CARBOHYD	821	821	N-LINKED (GLCNAC .)	(POTENTIAL)	-
SEQUENCE	861 AA;	97487 MM;	04DE2B4D64FD63A CRC64;		
Query Match	86.4%;	Score 2992.5;	DB 1;	Length 861;	
Best Local Similarity	86.7%;	Pred. No. 3.5e-224;			
Matches 569;	Conservative	32;	Mismatches 40;	Indels 15;	Gaps 8
QY	2	EKLWTVYVYGPVWMEATTTLFCASDAKAYOTEVHNVWATHACVPTDENPOEVLENTE	61		
DB	32	EKLWTVYVYGPVWKEATTTLFCASDAKAYOTEVHNVWATHACVPTDENPOEVLENTE	91		
QY	62	HNFWKNNWVEQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NAINF---INDSEG-	115		
DB	92	NFNWKKNDWVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLDGAVNTNSNNTSSSGE	151		
QY	116	-TMRGEIKNKSNPITTSIRBEVQKEYALFKLDVPIPDNNNTSKRLSCPTSVTQCP	174		
DB	152	MMMKGEIKNKSFINSTIRKVOKEYAFKLDIPIDNNTSTLTSCTSVTQCP	211		
QY	175	KISEPPIIHNCAPAGFAILKCNKTFNGKPPCKNVSVQCTHGIRPVSTOLLNGSLA	234		
DB	212	KVSEPIPIHNCAPAGFAILKCNKTFNGTGPCTNVSVQCTHGIRPVSTOLLNGSLA	271		
QY	235	EEBVVIRSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSIHI-GGRAFYTGEEIGD	292		
DB	272	EEBVVIRSANFTNAKTIIVOLQNSVEINCTRPNNNTRKSIIRIQGPRAFVYTIK-IGN	330		
QY	293	IRQAHCHISPRAKKNDTLKOIYIKLREQ-EKTIYFNHSSGGDEPITVHSPRCSEEFYC	351		
DB	331	MROAHCHISPRAKKNAITLKQISKREPGNNKTIIFKSSGGDEPITVHSPRCSEEFYC	390		
QY	352	NSTOLFNSTMNN---NTEGSSNTEG-NITPLPCRIKOIINMOVEGKMMVAPRIGQIRC	407		
DB	391	NSTOLFNSTMNSTWSTEGSNNTBSDTITLPCRKOIINMOVEGKMMAPRISGQIRC	450		
QY	408	SSNITGLILTRDGGINGENGTBIFPPGGDMKDMNRSEFYKTKVYKIEBPLGVAFTPKCKRV	467		
DB	451	SSNITGLILTRDGGNNNGNSEIFPPGGDMKDMNRSELYKTKVYKIEBPLGVAFTPKCKRV	510		
QY	468	VOREKRAVGAVPLGTGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQORM	527		
DB	511	VOREKRAVGIGALTEGFGAAGSTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQOHL	570		
QY	528	LQTVWVGKQIQARVIAVERLYGQOQLLGIVGSGSKLCCRAVPNNAWSMKSIDRINN	587		
DB	571	LQTVWVGKQIQARVIAVERLYKQOQLLGIVGSGSKLCTTAAVPNNAWSMKSISQIINN	630		
QY	588	MTWMEWEIREINITYSEIYTLIEESQNOQEKNEOELLEJDKWASLWNPEDITWMLWY	643		
DB	631	MTWMEWEIREINITYSLIHSLEEESQNOQEKNEOELLEJDKWASLWNPENITWMLWY	686		

[illegible]

Query Match	86.4%	Score 2992	DB 1	Length 852
Best Local Similarity	86.2%	Pred. No. 3,86-224		
Matches 558	Conservative 41	Mismatches 42	Indels 6	Gaps 4
QY 2 EKLWTVYGVGPVWKEATTTLFCASDAKAYOTEVHNWATHACVPTDPNPOEVLENTE	61			
DB 32 DKLWTVYGVGPVWKEATTTLFCASDAKAYOTEVHNWATHACVPTDPNPOELVGNANTE	91			
QY 62 HFNMMKNMMVBOEMOEDIIISLMDOSLPCVKLTPLCVTLNCKDVANATNTNDSSEGTMERGE	121			
DB 92 NFNMMKNMDMVEQMHEDIISLMDOSLPCVKLTPLCVTLNCHDPANATNATNSGKMMEGE	151			
QY 122 IKNCSEFNITTSIRDEVOKEVALPYKLDVPIIDNN--NTSYRLISDTSVITQACPISFE	179			
DB 152 MKNCSEFNITTSIRDKMKQEVAPLYKLDVPIIDNDKNTNRYLLSCNTSVITQACPITFE	211			
QY 180 PIPHYCAPAGFALLKNDKTFNKGKPCKNSTVQCTHGIRPVVSTOILLNGSLAEDEV	239			
DB 212 PIPHYCAPAGFALLKNNKKNKFNSTGPTNVSTVQCTHGIRPVVSTOILLNGSLAEDEV	271			
QY 240 IRSDFNFTNNAITIIIVOLKESVEINCTPRNNNTKRSIHIGPRAFFTTGEIIGDIRQACHN	299			
DB 272 IRSDFNFTNNAITIIIVOLKESVEINCTPRNNNTKRIITMGPRGVYTTQIIGDIRACHN	331			
QY 300 ISRAKMDTLKOIIVIKLEOPEKNTIVFNHSSGDPPEIIVMHSFNCEGEFFYCNSITOLPNS	359			
DB 332 LRSKXEMETLKOIIVTKLRVQPKNTIVFNRSGGDPEIIVMHSFNCGEFPFNTTOLPNS	391			
QY 360 TWNNNTSGSNTEBNT--ITLPCRKOIIMMOEYGVKAMYPPIRQIICSSNITGLLTR	418			
DB 392 TWYNTITG--NTEBNSPTLPCRKOIIMMOEYGVKAMYPPIRQIICSSNITGLLTR	450			
QY 419 DGGINE--NGTEIFRRPGGDMRDMRSEFYKYVVKIEPLDVAPTKCKRRVVOREKRAVG	476			
DB 451 DGGNNNETTDEIFRRPGGNNRDMRMRSELYKYVVKIEPLDVAPTKARRVVOREKRAVG	510			
QY 477 IGAVFLGFLGAAGSTMGASMTLLVQARLLLSGIVQOONNLRAIEAOQRMLOLTVWGIK	536			
DB 511 LGAFLFLGFLGAAGSTMGASLTLVQARLLLSGIVQOONNLRAIEAOQHLELTVWGIK	570			
QY 537 QLOARVLAVERYLGDQQLLGIGWCSGKLICTTAPVPAASWNSKSLDIRIWNNTWMEWRE	596			
DB 571 QLOARVLAVERYLKDQQLLGIGWCSGKLICTTAPVPAASWNSKSLDIWNTWMEWRE	630			
QY 597 IDNTSTSEIYTLIESONOQEKNEBELLELDKMASLMMNFDITNNIMLY	643			
DB 631 IDNTSTNLISLIEDSQIOQEKNEKLELLELDKMASLMMNFDITNNIMLY	677			

RESULT 8

ENV\_HY1B1 STANDARD; PRT; 856 AA.

AC P03375;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUN-1999 (Rel. 38, Last annotation update)

DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].

ENV.

Query Match	86.0%: Score 2978; DB 1; Length 856;	86.5%: Pred. No. 4.6e-223;	Matches 563; Conservative 35; Mismatches 43; Indels 10; Gaps 6
FT CARBOHYD	397	397	N-LINKED (GLCNAC . . .)
FT CARBOHYD	406	406	N-LINKED (GLCNAC . . .)
FT CARBOHYD	448	448	N-LINKED (GLCNAC . . .)
FT CARBOHYD	463	463	N-LINKED (GLCNAC . . .)
FT CARBOHYD	611	611	N-LINKED (GLCNAC . . .)
FT CARBOHYD	616	616	N-LINKED (GLCNAC . . .)
FT CARBOHYD	625	625	N-LINKED (GLCNAC . . .)
FT CARBOHYD	637	637	N-LINKED (GLCNAC . . .)
FT CARBOHYD	674	674	N-LINKED (GLCNAC . . .)
FT CARBOHYD	750	750	N-LINKED (GLCNAC . . .)
FT CARBOHYD	816	816	N-LINKED (GLCNAC . . .)
SEQUENCE	856 AA; 97224 MM; 08FBLAI931BB27 CRC64;		
Query Match	86.0%: Score 2978; DB 1; Length 856;		
Best Local Similarity	86.5%: Pred. No. 4.6e-223;		
Matches 563; Conservative 35; Mismatches 43; Indels 10; Gaps 6			
OY	2	EKLWVTVYGVGPVKEATNTTLFCASDAKAYOTEVHNWATHACVPTDENPOEVLLENTE	61
DB	32	EKLWVTVYGVGPVKEATNTTLFCASDAKAYOTEVHNWATHACVPTDENPOEVLLENTE	91
OY	62	HNWKNKNMVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCKDVANATNTDSEG--TWER	119
DB	92	NFNWKNQDMVQWQEDIIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGMIMEX	151
OY	120	GEINCKSFNTITSIRDEYQKYALFYKLDVPIPNNTNYSKLLISDTSVITQACXKISFE	179
DB	152	GEINCKSFNTISTSIRGKQKVEAFYFKLDIIPINDTTSYLTSCNTSVITQACKVSFE	211
OY	180	PIPIHYCAPAGFALILKNDKTFENKGPCKANSTVQCTGIRPVVSTOLLNGLSLAEVEV	239
DB	212	PIPIHYCAPAGFALILKCNKTKFNCTGIRPVVSTVQCTGIRPVVSTOLLNGLSLAEVEV	271
OY	240	IRSDNFTNNAKTIIVOLKESVEINCTRENNTRKSIHI--GPRGAFYTTGELIGDIRQAH	297
DB	272	IRSANFTDNNAKTIIVQLNQSVVEINCTRENNTRKSIIRGPRGAFYTTGK--IGNROAH	330
OY	298	CNIRPAKNDTLKQIVIKLRQF--BNKTIIVNHSAGDPETIVMHSFNCGEFFYCNSTOL	356
DB	331	CNIRPAKNNTLKQIDSKLRQFGNNKTIIPKQSSGSGPEIIVTHSFNCGGEFFYCNSTOL	390
OY	357	FNSTWPMN--NTBESNNNEG--NTITLPRIKQIINMWOEYKAMVAPIRGIRCSSNIT	412
DB	391	FNSTWPMSTWSTKSSNNEGSDTITLPRIKQIINMWOEYKAMVAPIRGIRCSSNIT	450
OY	413	GLTLTRDGLIENGTEIFRPGCGDMRDWRSEFYKVKVLEPLGVAETCKRRVVOREK	472
DB	451	GLTLTRDGSNNNSEIFRPGCGDMRDWRSELYKVKVLEPLGVAETKARRVVOREK	510
OY	473	RAVGIGAVFLGLGAAGSTMGAASTLTVQARLLISGIVQOONNLLRAIEAQORMLQTV	532
DB	511	RAVGIGALFLGLGAAGSTMGAASTLTVQAROLLSGVQOONNLLRAIEAQORMLQTV	570
OY	533	WGIQOLARVLAVERVYGLDQQLIGWCGSGKLICTTAVPMNASMNSKSLDIRIANNMTWE	592
DB	571	WGIQOLQRLIATVERVYGLDQQLIGWCGSGKLICTTAVPMNASMNSKSLDIRIANNMTWE	630
OY	593	WEREIDNTSEIYTLIEBSQNOEKNEQELLEDKWASLWMTDITNNLWY	643
DB	631	WDRFINNTSLIHSIEBSQNOEKNEQELLEDKWASLWMTDITNNLWY	681
RESULT 9			
ENV_HVIMN			
ID_ENV_HVIMN	STANDARD;	PRT;	856 AA.
AC	P05877;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-NOV-1988 (Rel. 09, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Envelope polypeptide Gp160 precursor [Contains: Exterior membrane glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].		
GN	ENV.		
OS	Human immunodeficiency virus type 1 (MN isolate) (HIV-1).		

[illegible]

Query Match	Best Local Similarity	86.0%;	Score 2978;	DB 1;	Length 856;
Matches 565;	Conservative 29;	Mismatches 43;	Indels 20;	Gaps 6	
FT CAROXYD 638 638	N-LINKED (GLCNAC...)	(POTENTIAL).			
FT CAROXYD 816 816	N-LINKED (GLCNAC...)	(POTENTIAL).			
SEQ SEQUENCE 856 AA;	97140 MM;	D197D80940BE732	CRG64;		
Query Match	Best Local Similarity	86.0%;	Score 2978;	DB 1;	Length 856;
Matches 565;	Conservative 29;	Mismatches 43;	Indels 20;	Gaps 6	
QY 2	EKLAVTVYGVVPEKAEATTTTFCASDAKAYDTEVHNWATACVPPDPNPOEVLNVT	61			
DB 31	EKLAVTVYGVVPEKAEATTTTFCASDAKAYDTEVHNWATACVPPDPNPOEVLNVT	90			
QY 62	HNHNKNNNNVBOOMODIISLMDSIKPCVKTLPLCAVLNCKXV-NATNTTN-----DS	113			
DB 91	HNHNKNNNNVBOOMODIISLMDSIKPCVKTLPLCAVLNCKXV-NATNTTN-----DS	150			
QY 114	EGTMRGEIKNCSFPIITTSIRDVEQKEYALFKLADVPIIDNNNTSYRLISCDPSVITQAC	173			
DB 151	EGTIRGEMKNCSPITTSIRDMQKEYALFKLADVPIIDNNNTSYRLISCDPSVITQAC	210			
QY 174	PKISFEPPIHYCAGAFALIKCNDKTFNGKGPCXNVSTVQCTHGIRPVYSTQLLNGSL	233			
DB 211	PKISFEPPIHYCAGAFALIKCNDKTFNGKGPCXNVSTVQCTHGIRPVYSTQLLNGSL	270			
QY 224	AAEEVVISDNTNNAKTIIVOLKASVEINCTRPNNNTRKSHIGEGRAPYTTGEIIGDI	293			
DB 271	AAEEVVISDNTNNAKTIIVOLKASVEINCTRPNNNTRKSHIGEGRAPYTTGEIIGDI	330			
QY 294	ROAHGNIRAKKNDTLKOIVILAREOPEKTI VPHNSSGSDPDEIWNHFNCGEFPYCN	353			
DB 331	ROAHGNIRAKKNDTLKOIVILAREOPEKTI VPHNSSGSDPDEIWNHFNCGEFPYCN	390			
QY 354	TQLENSYWN-----NTEGSNNTBGNITTLPCRKOIIMMOBVGKAMYAPPIRGIRCS	408			
DB 391	SPLFSTWNGNNTNNTGSGN-----NITLQCKIKOIIMMOBVGKAMYAPPIRGIRCS	446			
QY 409	SNITGLLITRPG--INENGTEIFRPGCGDNRDNRSEPKYKVKYLEPLGVAPTYCKR	466			
DB 447	SNITGLLITRPGKDTNDTEIFRPGCGDNRDNRSEPKYKVKYLEPLGVAPTYCKR	506			
QY 467	VVOREKRAVIGAVPLGFAGAGSTMGASMTLVTOARLLSGIVQOONNLRAIEAOR	526			
DB 507	VVOREKRA-AIGALPLGFAGAGSTMGASMTLVTOARLLSGIVQOONNLRAIEAOR	565			
QY 527	MLQLTWGIKOLOAVLAVERYLGDQQLLGIVGCSGKLICTTAVPNWASWNSKSLDRIN	586			
DB 566	MLQLTWGIKOLOAVLAVERYLGDQQLLGIVGCSGKLICTTAVPNWASWNSKSLDRIN	625			
QY 587	NNTWMEWERIDNTYSEIYTLIESONQOEKNEQELBELDKWASIMNWEDITNNLWY	643			
DB 626	NNTWMEWERIDNTYSEIYTLIESONQOEKNEQELBELDKWASIMNWEDITNNLWY	682			
RESULT 10					
ID ENV_HVLPV STANDARD; PRT; 856 AA.					
AC P03376;					
DT 21-JUL-1986 (Rel. 01, Created)					
DT 21-JUL-1986 (Rel. 01, Last sequence update)					
DT 15-JUL-1999 (Rel. 38, Last annotation update)					
DE Envelope polypeptide GP160 precursor [Contans: Exterior membrane					
GN glycoprotein (GP120); Transmembrane glycoprotein (GP41)].					
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).					
CC Virusess; Retroid virusess; Retroviridae; Lentivirus.					
CC NCBI_TaxID=11700;					
FN SEQUENCE FROM N.A.					
RX MEDLINE=8511157; PubMed=2982104;					
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,					
RA Capon D.J.,					
RT "Nucleic acid structure and expression of the human					
AIDS/Lymphadenopathy retrovirus."					

RL Nature 313:450-458(1985).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; K02083; AAB59873.1; -; ALT\_SEQ.  
 DR EMBL; X01762; CAA25903.1; ALT\_SEQ.  
 DR PIR; A03974; VCLJVL.  
 DR HIV; K02083; ENVSPV22.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Glycoprotein; Transmembrane;  
 Signal.

FT CHAIN 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 218 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97339 MW; 5FCD81DC3C1209B3 CRC64;

Query Match 85.9%; Score 2974; DB 1; Length 856;  
 Best Local Similarity 86.3%; Pred. No. 9,5e-223;  
 Matches 562; Conservative 34; Mismatches 45; Indels 10; Gaps 6;

QY 2 EKWVTVYGVVWVKEATTTTCASDARAYDTEVNWATACVPTDNDPEVLTENTE 61  
 DB 32 EKWVTVYGVVWVKEATTTTCASDARAYDTEVNWATACVPTDNDPEVLTENTE 91

QY 62 HFNMMKNVNEQNEQDIISLMDQSLKPCVLTCLNCKDVANNTINDSG--TMR 119  
 DB 92 NFNMMKNVNEQNEQDIISLMDQSLKPCVLTCLNCKDVANNTINDSG--TMR 151  
 QY 120 GEIKNGSFNTTTSIRBQVQKVALFYKLDVPIIDNNNTSRLSCDTSVTOACPKISFE 179  
 DB 152 GEIKNGSFNTTTSIRBQVQKVALFYKLDVPIIDNNNTSRLSCDTSVTOACPKISFE 211  
 QY 180 PIPIHCAFPALILKCNNDTFFNGKPCCKNVSTVQCCTHIGIRPVSTOLLNGSLAEV 239  
 DB 212 PIPIHCAFPALILKCNNDTFFNGKPCCKNVSTVQCCTHIGIRPVSTOLLNGSLAEV 271  
 QY 240 IRSDNFTNNAKTIIVQKSEVEINCTRPNNTRKSHI--GGRAPFTTGEITGDIRQAH 297  
 DB 272 IRSDNFTNNAKTIIVQKSEVEINCTRPNNTRKSHI--GGRAPFTTGEITGDIRQAH 330  
 QY 298 CNISRAKMNLTLOIYKLRQEF-EKTIYFNHSSGSDPEIVNHSFNGCEGFEYCNSTOL 356  
 DB 331 CNISRAKMNLTLOIYKLRQEF-EKTIYFNHSSGSDPEIVNHSFNGCEGFEYCNSTOL 390  
 QY 357 FNTGNN--NTGSGNNTG--NTITLPCRKOIINMVEGKAMYAPPIRGQIRCSNIT 412  
 DB 391 FNTGNN--NTGSGNNTG--NTITLPCRKOIINMVEGKAMYAPPIRGQIRCSNIT 450  
 QY 413 GLLLTDRGNGENGTETIFRPGGDMRDNMSEFYKRVKIEPLGVAFYKRRVQREK 472  
 DB 451 GLLLTDRGNGENGTETIFRPGGDMRDNMSEFYKRVKIEPLGVAFYKRRVQREK 510  
 QY 473 RANGIGVPLFGFGAAGSTMGASMTLVQARLLSGIVQOONNLRAIEAQRMLQTV 532  
 DB 511 RANGIGVPLFGFGAAGSTMGASMTLVQARLLSGIVQOONNLRAIEAQRMLQTV 570  
 QY 533 WGIKQDARVLAVERVYGDQQLIGWCSGKLTCTAVPNNASMSNKLDRINNNMTWME 592  
 DB 571 WGIKQDARVLAVERVYGDQQLIGWCSGKLTCTAVPNNASMSNKLDRINNNMTWME 630  
 QY 593 WERIDNYSIEITLIESONQOEKNEBELIEDKMASLNMWDITWMLY 643  
 DB 631 WERIDNYSIEITLIESONQOEKNEBELIEDKMASLNMWDITWMLY 681

RESULT 11  
 ENV\_HVLA2 STANDARD; PRT; 855 AA.  
 ID ENV\_HVLA2  
 AC P03378;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11685;  
 RN [1]  
 RP MEDLINE=8509453; PubMed=2578227;  
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,  
 RA Stempien M.W., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,  
 RA Levy J.A., Dina D., Luciw P.A.;  
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus  
 (ARV-2)".  
 RL Science 227:484-492(1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; K02007; AAB59882.1; -;

```

DR PIR; A03976; VCLJ2.
DR HIV; K02007; ENVSSF2.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
DR AIDS; Coat protein; Glycoprotein; Transmembrane;
Signal.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 29
FT CHAIN 30 509
FT CHAIN 510 855
FT DISULFID 53 73
FT DISULFID 118 208
FT DISULFID 125 199
FT DISULFID 130 155
FT DISULFID 221 250
FT DISULFID 231 242
FT DISULFID 299 333
FT DISULFID 380 442
FT DISULFID 387 415
FT CARBOHYD 87 89
FT CARBOHYD 129 129
FT CARBOHYD 140 140
FT CARBOHYD 154 154
FT CARBOHYD 158 158
FT CARBOHYD 184 184
FT CARBOHYD 190 190
FT CARBOHYD 200 200
FT CARBOHYD 223 233
FT CARBOHYD 244 244
FT CARBOHYD 265 265
FT CARBOHYD 279 279
FT CARBOHYD 292 292
FT CARBOHYD 298 298
FT CARBOHYD 304 304
FT CARBOHYD 334 334
FT CARBOHYD 341 341
FT CARBOHYD 358 358
FT CARBOHYD 364 364
FT CARBOHYD 388 388
FT CARBOHYD 394 394
FT CARBOHYD 400 400
FT CARBOHYD 408 408
FT CARBOHYD 445 445
FT CARBOHYD 458 458
FT CARBOHYD 461 461
FT CARBOHYD 610 610
FT CARBOHYD 615 615
FT CARBOHYD 624 624
FT CARBOHYD 636 636
FT CARBOHYD 815 815
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC1A2 CRC64;

Query Match 85.8%; Score 2972; DB 1; Length 855;
Beet Local Similarity 86.2%; Pred. No. 1.4e-222;
Matches 563; Conservative 32; Mismatches 44; Indels 14; Gaps 8;

2 EKLWTVVYGVVWKEATTTLFCASDARAYDTEVNNWATAHCVPTDNPQVLEANTE 61
31 EKLWTVVYGVVWKEATTTLFCASDARAYDTEVNNWATAHCVPTDNPQVLEANTE 90
62 HFNMMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTME-RG 120
91 NFNMMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTME-RG 150
121 EIKNCSFNITTSIRDEVOKEVALPYKLDVPIDN-----NNTSYRLISCDTSVITQACP 175
151 EIKNCSFNITTSIRDEVOKEVALPYKLDVPIDN-----NNTSYRLISCDTSVITQACP 210
176 ISFERIPHYCAPAFALIKCKDKTPNGKPKANSTVQCTGTGIRPVVSTOLLNGSLAE 235
211 VSFEPPIHYCTPAGFAILKCKNNKTFNGKPGCTNVTVOCTGIRPIVSTOLLNGSLAE 270

```

```

QY 236 EEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTKRSIHIGCPRAFTYTGEIIGDIRQ 295
DB 271 EEEVIRSDNFTNNAKTIIVQLKESVAINCTRPNNNTKRSIYIGCPRAFTTGRIGDIRK 330
QY 296 AHCHNSRPAKMDTLKQIVIKAREOF-EKKTIVPNHSSGGDEIYHSHNCGEPFYCNST 354
DB 331 AHCHNSRPAKMDTLKQIVIKAREOF-EKKTIVPNHSSGGDEIYHSHNCGEPFYCNST 390
QY 355 QLFNSTWN-NNTGSSNNTGNN-TITLPCRIOIIMMOEVCAMAYAPPIRQIRCSSNIT 412
DB 391 QLFNSTWN-NNTGSSNNTGNN-TITLPCRIOIIMMOEVCAMAYAPPIRQIRCSSNIT 447
QY 413 GLLLTRDGIN-ENGTEIFRPGGDMRDMRSSEFYKRVKXI EPLGVAPTCKRRVQRE 471
DB 448 GLLLTRDGIN-ENGTEIFRPGGDMRDMRSSEFYKRVKXI EPLGVAPTCKRRVQRE 507
QY 472 KRAVGI-GAVLGLFGAAGSTMGASMTLTQARLLSGIVQOONNLRALEAOQRLQL 530
DB 508 KRAVGI-GAVLGLFGAAGSTMGASMTLTQARLLSGIVQOONNLRALEAOQRLQL 567
QY 531 TWGIKQIQAARVLAVERIAGPQOLLGIVGSGSKLICTTAVPMNASWSKSLDRINNMNTM 590
DB 568 TWGIKQIQAARVLAVERIAGPQOLLGIVGSGSKLICTTAVPMNASWSKSLDRINNMNTM 627
QY 591 MEMEREIDNVTSEIYTLIESQNOEKNEQELLEDKVASIMNPFITNMLMY 643
DB 628 MEMEREIDNVTSEIYTLIESQNOEKNEQELLEDKVASIMNPFITNMLMY 680

RESULT 12
ENV_HV1H2 STANDARD; PRT; 856 AA.
AC P04578; 009779;
ID 13-AUG-1987 (Rel. 05. Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP120 precursor [Contains: Exterior membrane
GN glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; Pubmed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RT "Complete nucleotide sequences of functional clones of the AIDS
RL virus."
RN AIDS Res. Hum. Retroviruses 3:57-69 (1987).
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RL Gallo R.C., Wong-Staal F.;
RC Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF033819; AAC82596.1; -
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9W; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.

```



DR PDB: 1K33; 10-OCT-01.  
 DR PDB: 1K34; 10-OCT-01.  
 DR HTV; K03455; ENVSHK2.  
 DR InterPro: IPR00328; Env\_GP41.  
 DR InterPro: IPR00777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 DR AIDS: Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;  
 3D-structure.  
 KM SIGNAL 1 30  
 FT CHAIN 31 511  
 FT CHAIN 512 856  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT CARBOHYD 88 418  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 356 356  
 FT CARBOHYD 386 386  
 FT CARBOHYD 392 392  
 FT CARBOHYD 397 397  
 FT CARBOHYD 406 406  
 FT CARBOHYD 448 448  
 FT CARBOHYD 463 463  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 624 624  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 750 750  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97212 MW; 6FA16AF85107FEO CRC64;  
 Query Match 85.6%; Score 2964; DB 1; Length 856;  
 Best local similarity 85.0%; Pred. No. 5.7e-22;  
 Matches 560; Conservative 36; Mismatches 45; Indels 10; Gaps 6;

QY 240 IRSDNFTNNAKTIIVOLKESEVEINCRPNNTKRSIHI--GPGAFYTTGEIGDIRQAH 297  
 DB 272 IRSVNFTDNAKTIIVOLNTSVEINCRPNNTKRIIRIQGPGAFYTTGEIGDIRQAH 330  
 QY 298 CNISRAKMNNDLKOIVYIKREOF-ENKTIVENHSSGGDEPIVMSFNCSEGFYCNSTOL 356  
 DB 331 CNISRAKMNNDLKOIVYIKREOF-ENKTIVENHSSGGDEPIVMSFNCSEGFYCNSTOL 390  
 QY 357 FNSGTNN--NTEGSSNTEG--NTITLPCRKOIINMVEYKAMYPPIGQIRCSNIT 412  
 DB 391 FNSGTNN--NTEGSSNTEG--NTITLPCRKOIINMVEYKAMYPPIGQIRCSNIT 450  
 QY 413 GLILTRGGINNGEITFRPGGDMRNMSSEFYKYVYKIEPLGVAFTCKRRVYREK 472  
 DB 451 GLILTRGGSNNSEIFRPGGDMRNMSSEFYKYVYKIEPLGVAFTCKRRVYREK 510  
 QY 473 RAVGIGAVPLGFGAAGSTWGAASMTLTVQARLLSGIVQOONNLLRAIEAQPMQLTV 532  
 DB 511 RAVGIGAVPLGFGAAGSTWGAASMTLTVQARLLSGIVQOONNLLRAIEAQPMQLTV 570  
 QY 533 WGIKOLARVLAVERVYIGDQDLGIVGCGSKLICCTAVPNASMSKSLRINNNMTWE 592  
 DB 571 WGIKOLARVLAVERVYIGDQDLGIVGCGSKLICCTAVPNASMSKSLRINNNMTWE 630  
 QY 593 WEREIDVTSEITVIESONOEKNEOBLELDKMASLNWPFDTWMLY 643  
 DB 631 WEREIDVTSEITVIESONOEKNEOBLELDKMASLNWPFDTWMLY 681  
 RESULT 13  
 ENV HV1W1 STANDARD; PRT; 856 AA.  
 ID ENV HV1W1 STANDARD; PRT; 856 AA.  
 AC P31872;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.  
 OX NCBI\_Taxid=31678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=66218077; PubMed=2423250.  
 RA Starck B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;  
 RT "Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS."  
 RT AIDS.  
 RL Cell 45:637-648(1986).  
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEPARATEDLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC PIR: A24774; VCLJW.  
 DR PDB: 1LBO; 04-DEC-02.  
 DR PDB: 1LCK; 04-DEC-02.  
 DR InterPro: IPR00328; Env\_GP41.  
 DR InterPro: IPR00777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 DR AIDS: Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;  
 3D-structure.  
 KM SIGNAL 1 29  
 FT CHAIN 30 510  
 FT CHAIN 511 856  
 FT DISULFID 53 73  
 FT DISULFID 118 205  
 FT DISULFID 125 196  
 FT DISULFID 130 152  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 TRANSMEMBRANE GLYCOPROTEIN.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.



FT	DISULFID	396	330	BY SIMILARITY.
FT	DISULFID	276	444	BY SIMILARITY.
FT	DISULFID	383	417	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	151	151	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	155	155	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	183	183	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	331	331	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	354	354	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	390	390	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	394	394	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	404	404	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	447	447	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	459	459	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. . . ) (POTENTIAL).
50	SEQUENCE	856 AA;	97526 MW;	DB68D1B49C404DB9 CR664;

Query Match	85.5%	Score 2358.5	DB 1	Length 856
Best Local Similarity	86.1%	Pred. No. 1.5e-221		
Matches	564	Conservative	30	Mismatches 46; Indels 15; Gaps 7
QY	1	VEKLVMTYYGYGVPMKEATITTLFCASDAKADYTEVHANNWATHACYPTDBNPOEVLNENT	60	
Db	30	VEQLMVTYYGVPMKEATITTLFCASDAKADYTEHAKWATHACVPTMNPDEVLNENT	89	
QY	61	EHPFMKKNMVEOMQEDIIISLMDQSLKPCVKLTPLCVTLNCQDVNATNTNSEGMERG	120	
Db	90	ENFMKKNMVEOMQHEDIISLMDQSLKPCVKLTPLCVTLNCDKXITMENT--IIGGG	147	
QY	121	EIKNCSFNITTSIRDEVOKEVALFYKLDVVPIDNN---TSYRLISDTSVITQACBK	175	
Db	148	EVKNCSFNITTSIRKVKHEVALFYKLDVVPIDKSNNDSTYRYRLIHGNTSVITQACSK	207	
QY	176	ISPEPIPHYCAPAGFALIKCKDKTFENGKPGPKANSTVOCTGIRPVYSTOLLNGSLAE	235	
Db	208	VSPEPIPHYCAPAGFALIKCKDKTFNGTGPSTNASTVOCTGIRPVYSTOLLNGSLAE	267	
QY	236	EEVVISRDNFTNNAKTIIVOLKESVEINCTRENNT-RKSIHIGGRAFYTTGBIIGDIR	294	
Db	268	EEIVIRSENFIDNAKTIIVHLNESVEINCTRENNVRRRHIIIGGRAFY-TGEIRGNIR	326	
QY	295	QAHNCISAKKNKDTLKOIVILRBEFNKTIIVHNSGCDPEIIVHNSFRCBEEFFYCNST	354	
Db	327	QAHNCISAKKNKTKQIVELRBEFNKTIIVHNSGSDPEIIVHNSFRCGGEFFYCNST	386	
QY	355	QLFNSTVN--NNTGSGNNTS--GVTITLPCRKOIIMMOWEVRGAMYAPIRGOIRGCS	409	
Db	387	QLFNSTVNVTGISTEGNNTTEBNDTITLPCRKOIIMMOWEVRGAMYAPIRGOIRGCS	446	
QY	410	NITGILLTRDGINENGTEIFRPGGDMRDWNRSEFYKTKVYKIEPLGYAPFKCRKRYVQ	469	
Db	447	NITGILLTRDGNSSRREIFRPGGDMRDWNRSELVYKTKVYKIEPLGYAPFKARRRYVQ	506	
QY	470	REKRAVG-IGAVPLGFLGAAGSTWGAAMTLTVQARLLSGIVQOQNNTLRAIEAQOHL	528	
Db	507	REKRAVGAIIGAVPLGFLGAAGSTWGAASLTTLVQAROLLSGIVQOQNNTLRAIEAQOHL	566	
QY	529	QLTWGIRQLQARVLAVERYLGDQDLGIWGSGLGICCTAIVPNAASNKSLEDBRIWMNN	588	

[illegible]

[illegible]

FT CARBOHYD 812 812 N-LINKED (GLCMAC. . .) (POTENTIAL).  
SQ SEQUENCE 852 AA; 96663 MW; EB7BBF8D23C9910D CRC64;

Query Match 85.3%; Score 2954.5; DB 1; Length 852;  
Best Local Similarity 85.6%; Pred. No. 3.1e-221;  
Matches 555; Conservative 33; Mismatches 53; Indels 7; Gaps 5;

QY 2 EKLAVTVVYGVPMKEATTTLLFCASDPAKYDTEVHNWMAHACVPTDPNPOEVLENTYE 61  
DB 31 EKLAVTVVYGVPMKEATTTLLFCASDPAKYDTEVHNWMAHACVPTDPNPOEVLENTYE 90  
QY 62 HENMKNNMVEQMOEDIIISLMDQSLKPCYKLTPLCVTLNCKDV--NATNTNDSGEMER 119  
DB 91 HENMKNNMVEQMOEDIIISLMDQSLKPCYKLTPLCVTLNCKDV--NATNTNDSGEMER 150  
QY 120 GEIKKCSFNTITSIDEXEYALFYKLDVPI--DNNTSYRLISCDTSVITQACPKIS 177  
DB 151 GEIKKCSFNTITSIDEXEYALFYKLDVPI--DNNTSYRLISCDTSVITQACPKIS 210  
QY 178 FEPIPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGSLAEE 237  
DB 211 FEPIPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGSLAEE 270  
QY 238 VIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRAH 297  
DB 271 VIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRAH 330  
QY 298 CNISRAKNNDTLKOYVILKREOFENKTYFENHSSGGDEPEIVMHSFNCBGEFFYCSTOLF 357  
DB 331 CNISRAKNNDTLKOYVILKREOFENKTYFENHSSGGDEPEIVMHSFNCBGEFFYCSTOLF 390  
QY 358 NSTANNTEGNNTEGNTTTLPCRIKQIINMVEGKAMYAPPIRGQIRCSNITGLLLT 417  
DB 391 NSTANNTEGNNTEGNTTTLPCRIKQIINMVEGKAMYAPPIRGQIRCSNITGLLLT 449  
QY 418 RDGINENG-TEIFRPGGDMRDNRSEFEYKTKVYKIEPLGVAPTKCKRRVVQREKRAVG 476  
DB 450 RDGINENG-TEIFRPGGDMRDNRSEFEYKTKVYKIEPLGVAPTKCKRRVVQREKRAVG 509  
QY 477 -IGAVFLGFLGAAGSTMGAASMTLVQARLLSGIVQOONNLLRAIEAQOHLQLTWGI 535  
DB 510 -IGAVFLGFLGAAGSTMGAASMTLVQARLLSGIVQOONNLLRAIEAQOHLQLTWGI 569  
QY 536 KOLQARVLAVERYLADQDLGIMGCSGKLICTTAVPANASMSKSLDRIMNNMTWEMER 595  
DB 570 KOLQARVLAVERYLADQDLGIMGCSGKLICTTAVPANASMSKSLDRIMNNMTWEMER 629  
QY 596 EIDNTYSRIYTLIESQNOEKNQOELLELDKMASIMNFDITNMLMY 643  
DB 630 EIDNTYSRIYTLIESQNOEKNQOELLELDKMASIMNFDITNMLMY 677

Search completed: December 12, 2003, 12:30:46  
Job time : 11.7921 secs

**THIS PAGE BLANK (US 710)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:26:14 ; Search time 33.0726 Seconds  
(without alignments)  
5017.086 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYVYGVPMKKAAT.....ELDKWASLWMDITNLMWY 643

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

1 number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3424	98.9	847	15	Q75760 human immun
2	3116.5	90.0	848	15	Q74999 human immun
3	3086.5	89.2	852	15	Q92761 human immun
4	3075.5	88.8	854	15	O40222 human immun
5	3073.5	88.8	854	15	Q78225 human immun
6	3069.5	88.7	855	15	Q03805 human immun
7	3068	88.6	853	15	Q03811 human immun
8	3063.5	88.5	846	15	Q9PXW7 human immun
9	3059.5	88.4	838	15	Q03806 human immun
10	3059	88.4	851	15	O56110 human immun
11	3048	88.0	863	15	Q77989 human immun
12	3045.5	88.0	868	15	Q9WJW5 human immun
13	3044	87.9	854	15	O92762 human immun
14	3039.5	87.8	860	15	Q9YP50 human immun
15	3039	87.8	861	15	Q9WJVS human immun
16	3027	87.4	847	15	Q8AR21 human immun

17	3027	87.4	849	15	Q77368 human immun
18	3027	87.4	849	15	Q8Q851 human immun
19	3022.5	87.3	843	15	Q70008 human immun
20	3021	87.3	855	15	Q8UL64 human immun
21	3021	87.3	863	15	Q42031 human immun
22	3019	87.2	851	15	O56562 human immun
23	3018	87.2	863	15	Q9WJW8 human immun
24	3017.5	87.2	860	15	Q9YP43 human immun
25	3017	87.1	843	15	Q70150 human immun
26	3015	87.1	859	15	Q8Q850 human immun
27	3013	87.0	853	15	Q86108 human immun
28	3010	86.9	863	15	Q9WJU4 human immun
29	3009.5	86.9	841	15	Q9QXJ1 human immun
30	3008.5	86.9	843	15	Q9QXK8 human immun
31	3007.5	86.9	852	15	O41883 human immun
32	3007	86.9	853	15	Q9J022 human immun
33	3003.5	86.8	856	15	Q72993 human immun
34	3002.5	86.7	846	15	Q8UL63 human immun
35	3002.5	86.7	854	15	Q92875 human immun
36	3001	86.7	861	15	Q9E527 human immun
37	3000.5	86.7	856	15	O71270 human immun
38	3000	86.7	851	15	Q8G852 human immun
39	3000	86.7	861	15	Q9E523 human immun
40	2999	86.6	854	15	O56566 human immun
41	2999	86.6	855	15	Q8UL62 human immun
42	2998	86.6	859	15	O71260 human immun
43	2997.5	86.6	840	15	Q8Q2X4 human immun
44	2996	86.5	861	15	O56109 human immun
45	2995.5	86.5	850	15	Q70003 human immun

## ALIGNMENTS

RESULT 1

ID Q75760 PRELIMINARY; PRT; 847 AA.

AC Q75760;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_taxid=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=87206194; PubMed=3646751;

RA Koyanagi Y., Miles S., Mitsuhashi R.T., Merrill J.E., Vinters H.V., Chen I.S.;

RA "Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms.";

RT Science 236:819-822(1987).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=91043044; PubMed=2172833;

RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Zack J.A., Chen I.S.;

RA "HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain.";

RT Nature 348:69-73(1990).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=92092169; PubMed=1684385;

RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;

RA "HIV-1 env sequence variation in brain tissue of patients with AIDS-related neurologic disease.";

RT J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).

RL [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN=JREFL;  
 RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.,  
 RA Kojanagi Y., Namazie A., Zhao J., Diagne A., Jiler K.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63632; AAB05604.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00517; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 847 AA; 96160 MW; 022D5F24B04FB29F CRC64;

Query Match 98.9%; Score 3424; DB 15; Length 847;  
 Best Local Similarity 99.2%; Pred. No. 1.1e-276;  
 Matches 638; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 VEKLMVTYYGVVPWKEATTTTLFCASDAKAYDEVHNVMTACVPTDPNQEVLLENVT 60  
 30 VEKLMVTYYGVVPWKEATTTTLFCASDAKAYDEVHNVMTACVPTDPNQEVLLENVT 89  
 QY 61 EHFMMKNNMVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 120  
 90 EHFMMKNNMVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 149  
 QY 121 EIKKCSFNITTSIRDEVQKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 180  
 150 EIKKCSFNITTSIRDEVQKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 209  
 QY 181 IPIHYCAPAGFALLKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 240  
 210 IPIHYCAPAGFALLKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 269  
 QY 241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHCHN 300  
 270 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHCHN 329  
 QY 301 SRAKMDTLKOIVIKLREOFENKTVFNHSSGSDPEIVNHSFNCGEFFYCNSTOLFNS 360  
 330 SRAKMDTLKOIVIKLREOFENKTVFNHSSGSDPEIVNHSFNCGEFFYCNSTOLFNS 389  
 QY 361 WNNNTGSGNNTGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTRDG 420  
 390 WNNNTGSGNNTGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTRDG 449  
 QY 421 GINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAPTKAKRRVQREKRAVGIGAV 480  
 450 GINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAPTKAKRRVQREKRAVGIGAV 509  
 QY 481 FLGFLGAAGSTMGASMTLTVQARLLLSGIYVOQONNLRAIEAOQRMQLTVMGIKOLA 540  
 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIYVOQONNLRAIEAOQRMQLTVMGIKOLA 569  
 QY 541 RVLAVERYLGDQQLIGWCSGKLICTAVPWNASWSKSLDRINNNMTWMEWEEREIDNY 600  
 570 RVLAVERYLGDQQLIGWCSGKLICTAVPWNASWSKSLDRINNNMTWMEWEEREIDNY 629  
 QY 601 TSEIYTLIEESQNOEKNEOELELDKNASLNNWDTINWLMY 643  
 630 TSEIYTLIEESQNOEKNEOELELDKNASLNNWDTINWLMY 672

RESULT 2  
 ID Q74999 PRELIMINARY; PRT; 848 AA.  
 AC Q74999;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Envelope glycoprotein gp160.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JR-CSF;  
 RX MEDLINE=97063128; Pubmed=8906996;  
 RA Klasse P.J., Boyd M.T., Weiss R.A., Schulz T.F.;  
 RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing  
 RT variant of HIV type 1 JR-CSF that infects a range of T cell lines";  
 RL AIDS Res. Hum. Retroviruses 12:347-350(1996).  
 DR EMBL; U45960; AAB60591.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 848 AA; 96433 MW; B240019C8737BC3 CRC64;

Query Match 90.0%; Score 3116.5; DB 15; Length 848;  
 Best Local Similarity 90.5%; Pred. No. 1.4e-245;  
 Matches 584; Conservative 27; Mismatches 31; Indels 3; Gaps 3;

QY 1 VEKLMVTYYGVVPWKEATTTTLFCASDAKAYDEVHNVMTACVPTDPNQEVLLENVT 60  
 30 VEKLMVTYYGVVPWKEATTTTLFCASDAKAYDEVHNVMTACVPTDPNQEVLLENVT 89  
 QY 61 EHFMMKNNMVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 120  
 90 EHFMMKNNMVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 149  
 QY 121 EIKKCSFNITTSIRDEVQKEALFYKLDVVPIDN-NTSYRLISCDTSVITQACPKISFE 179  
 150 EIKKCSFNITTSIRDEVQKEALFYKLDVVPIDSKNTKRLINCNTSVITQACPKISFE 209  
 QY 180 PIPHYCAPAGFALLKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 239  
 210 PIPHYCAPAGFALLKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 269  
 QY 240 IRSNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHCHN 299  
 270 IRSNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHCHN 329  
 QY 300 ISRAKMDTLKOIVIKLREOFENKTVFNHSSGSDPEIVNHSFNCGEFFYCNSTOLFNS 359  
 330 ISRAKMDTLKOIVIKLREOFENKTVFNHSSGSDPEIVNHSFNCGEFFYCNSTOLFNS 389  
 QY 360 TWNNNTGSGNNTG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTR 418  
 390 TW-NDTEKSSGTEGDTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTR 448  
 QY 419 DGINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAPTKAKRRVQREKRAVGIG 478  
 449 DGINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAPTKAKRRVQREKRAVGIG 508  
 QY 479 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIYVOQONNLRAIEAOQRMQLTVMGIKOLA 538  
 509 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIYVOQONNLRAIEAOQRMQLTVMGIKOLA 568  
 QY 539 QARVLAVERYLGDQQLIGWCSGKLICTAVPWNASWSKSLDRINNNMTWMEWEEREID 598  
 569 QARVLAVERYLGDQQLIGWCSGKLICTAVPWNASWSKSLDRINNNMTWMEWEEREID 628  
 QY 599 NYTSEIYTLIEESQNOEKNEOELELDKNASLNNWDTINWLMY 643  
 629 NYTSEIYTLIEESQNOEKNEOELELDKNASLNNWDTINWLMY 673

RESULT 3	ID	092761	PRELIMINARY;	PTT;	852 AA.
AC	092761;				
DT	01-NOV-1998 (TRMBLrel. 08, Created)				
DT	01-NOV-1998 (TRMBLrel. 08, last sequence update)				
DT	01-MAR-2003 (TRMBLrel. 23, last annotation update)				
DE	Envelope glycoprotein.				
GN	ENV.				
OS	Human immunodeficiency virus 1.				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxId=11676;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SFMS2.9;				
RX	MEDLINE=98178716; PubMed=9519894;				
RA	McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K., Shippey W.H.;				
P	"Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in San Francisco Men's Health Study participants."				
DR	AIDS Res. Hum. Retroviruses 14:329-337(1998).				
DR	EMBL; AF025750; AAC40588.1; -				
DR	InterPro: IPR000328; Env.GP41.				
DR	InterPro: IPR000177; GAP_Chidogenase.				
DR	InterPro: IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
DR	PROSITE: PS00071; GAPDH; 1.				
KW	AIDS; Coat protein; Envelope glycoprotein; Polyprotein; Transmembrane.				
SO	SEQUENCE 852 AA; 96545 MW; 3C7780DB0611E617 CRC64;				

Query Match	89.2%	Score 3086.5	DB 15	Length 852
Best Local Similarity	89.5%	Pred. No. 4.1e-243		
Matches	579	Conservative	25	Mismatches 36; Indels 7; Gaps 5

  

Qy	3	KLMVTVVYGVVPWKEATITLLFCASADAKADTEVHNWMACTACVPTPNPQEVLENVTEH	62
Db	32	QLMVTVVYGVVPWKATITLLFCASDAKADTEVHNWMACTACVPTPNPQEVLVGVNTEH	91
Qy	63	FNMKNNVVEQMEODIISLMDSLKPCVKLTPLCTVLNCDV-NATNTNDSGTMERGE	121
Db	92	FNIMKNNVVEQMEHDIISLMDSLKPCVKLTPLCTVLNCDTLDNRANATNTSSSGEMMERGE	151
Qy	122	IKNGSFNTTISRDEQKAYALFYKLDVVPID-NNTSYRLISGTSVYTAQCPKISFE	179
Dy	152	IKNGSFNTTISRDKQKEYALFYKLDVVPIDNNNTSYRLISGTSVYTAQCPKVSFE	211
Qy	180	PIPIHYCAPAGFALLKCDKTEGNGKPCKNVSTVQCTHGRIPVSTOLLNGSLAEVY	239
Dy	212	PIPIHYCAPAGFALLKCNKTEFNNGSQCCTNVSTVQCTHGRIPVSTOLLNGSLAEVY	271
Qy	240	IRSDFTNNAKTIIVQLKESVEINCTRPNNNTKRSIHIGRAFYTTGSLIGDIRAHCN	289
Dy	272	IRSEITNNAKTIIVQLKEPVEINCTRPNNNTKRSIHIGRAFYTTGSLIGDIRAHCN	331
Qy	300	ISPAKNDTLKOIVIKREFQENKTIIVFHSQSGDEIVMHSNCGSEFPYCNSTOLFNS	359
Dy	332	ISPAKNDTLKRIQAOLKRFQNKTIIVFQSGDEIVYTHSNCGSEFPYCNASOLFNS	391
Qy	360	TWNNTEGSSNTEGN-TITLPCRIKQIINMVEQKAMYPARIRGOIRCSSNITGLLT	417
Dy	392	TW-NDTERSSNTEGNDTITLPCRIKQIINMVEQKAMYPARIRGOIKCASNITGLLT	450
Qy	418	RDGGINERG-NEIIRPGGDMRDMNRSEYKTKVYKIEPLGVAPYTCRKRYYVOREGAAG	476
Dy	451	RDGGNNNTGTETTFRPGGDMRDMNRSELYKTKVYKIEPLGAPYPAKRRYYVOREKRAAG	510
Qy	477	IGAVFLGFLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAQORMLQITWYGK	536
Dy	511	IGAVFLGFLGAAGSTMGASVTLTVOARLLLSGIYQOQNNLLRAIEAQOHLQITWYGK	570
Qy	537	QLQARVLAVERYLDQOQLIGTGCSEGLICTTAVPWNASWSNKSIDRINNNTMEWERE	596

D3 571 QLOARVLAVERTILRDQOLIGMGSGKLCITTPYPMNTSMNSKSLDTIIMNNTIMQWERB 630

D4 597 IIDNYTSEITVLIEESQNOEKKEQLELLELDKMASIMNFPDITNNLWY 643

D5 631 IDNYTSLIYTLIEESQNOEKKEQLELLELDKMASIMNFPDITNNLWY 677

ID	Accession	Source	Length
Q40222		PRELIMINARY;	PRT; 854 AA.
DT	Q40222		
DT	01-JAN-1998 (TrEMBLrel. 05, Created)		
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)		
DE	Env. Polyprotein.		
GN	ENV.		
OS	Human immunodeficiency virus 1.		
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RY	[1]		
RY	SEQUENCE FROM N.A.		
RC	STRAIN=AD8;		
RX	MEDLINE=9643129; PubMed=8835195;		
RA	Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,		
RA	Martin M.A., Peden K.W.,		
RT	"Construction and characterization of a stable full-length macrophage		
RT	tropic HIV type 1 molecular clone that directs the production of high		
RL	titers of progeny virions.";		
RL	AIDS Res. Hum. Retroviruses 12:191-194(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AD8;		
RA	Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,		
RA	Martin M.A., Peden K.W.C.;		
RA	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF004394; AAB64170.1; "		
DR	InterPro; IPR000328; Env GP41.		
DR	InterPro; IPR000777; GP120.		
DR	Pfam; PF00516; GP120; 1.		
DR	Pfam; PF00517; GP41; 1.		
KW	AD8; Coat protein; glycoprotein, Polyprotein, Transmembrane.		
QO	SEQUENCE 854 AA; 97291 MW; 06C45B69103C6C12 CRC64;		

Query Match	88.8%	Score 3075.5	DB 15	Length 854
Best Local Similarity	89.4%	Pred. No. 3.3e-242		
Matches 583	Conservative 19	Mismatches 39	Indels 11	Gaps 7
Qy	1	VEKLVATVYVGVPMKEATTTLP	PCASDAKADVTEVHNVMATHACVPTDPNPOEVL	ENT 60
Dp	30	VEKLVATVYVGVPMKEATTTLP	CASDAKADVTEVHNVMATHACVPTDPNPOEVL	ENT 89
Qy	61	EHHNNKKNNMVQOMEDIIISL	DSLKRCVLTPLCVTLTLCNKDV	-NATNTTNSSEGIMER 119
Dp	90	ENHNKKNNMVQOMEDIIISL	DSLKRCVLTPLCVTLCTLNTD	LRNTNTNNINSSEGIM -R 147
Qy	120	GEIKNCSPFITTSIRDEYOEK	EALPYKLDVPIINNNTS	YRLISCDTSVITQACPKISFE 179
Dp	148	GEIKNCSPFITTSIRDKYKDY	ALFYRLDVPIDINDTSTR	LINCNTSTITQACPKISFE 207
Qy	180	PIBIHYCABAGFAILKCKNDK	TENGKGPCANSTVQCTHGIR	PPVSTOLLNGSLAEEBV 239
Dp	208	PIBIHYCTBAGFAILKCKDK	KFKNGTGPCANSTVQCTHGIR	PPVSTOLLNGSLAEEBV 267
Qy	240	IRSDNFTNNAKIIIVQLKESV	EINCTRPNNNTKRSIHIGGR	AFYTTGELIGDIRQACHN 299
Dp	268	IRSNFTNDAKNIIVQLKESV	EINCTRPNNNTKRSIHIGGR	AFYTTGELIGDIRQACHN 327
Qy	300	ISRAKKNDIILKQIVIKLREOF	-ENKTIIPFNHSSGSGDEI	VMHSPNCGEFPYCNSITOLFV 358
Dp	328	ISRTKKNNTIANDIARKLEOF	GNNKTIIVFNHSSGSGDEI	VMHSPNCGEFPYCNSITOLFV 387
Qy	359	STWNN-----NTEGSNNTEGN	-TITLPCR	IKQIIMMOBVGKAMVAPIRGQIRCSSNITG 413

Db 388 STNFMNGTWNLTQSGTEGNDITLPCRIKQIIMMVGKAMAPPIRGQIRCSNITG 447

Qy 414 LLTRDGINENG-TEIRPRGGMDRMRSEFYKVKVKEIPLGVAATKCKRRVQREK 472

Db 448 LILTRDGGNNNNNDTEFRPGGDMRDNRSELYKVKVKEIPLGVAATKCKRRVQREK 507

Qy 473 RAVG-IGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOQRLQTLTWG 531

Db 508 RAVGTIGAMFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOQRLQTLTWG 567

Qy 532 VMGIKQIQAARVLAVERLYGDOQLLGIMGCGSKLICCTAVPWNASWSKSLDRIMNNMTWM 591

Db 568 VMGIKQIQAARVLAVERLYGDOQLLGIMGCGSKLICCTAVPWNASWSKSLDRIMNNMTWM 627

Qy 592 EMEEINNTSEIYTLIEESONOEKNEQELBELDKKASIMNPDITNMLMY 643

Db 628 EMEEIDNYTLGTLIEESONOEKNEQELBELDKKASIMNPDITNMLMY 679

Query Match 88.8%; Score 3073.5; DB 15; Length 854;  
Best Local Similarity 88.6%; Pred. No. 4.7e-242;  
Matches 575; Conservative 24; Mismatches 43; Indels 7; Gaps 2;

Qy 2 EKLWTVYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61

Db 31 EKLWTVYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90

Qy 62 HFNMMKNMVEOMOEEDISLMDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115

Db 91 NFNMMKNMVEOMOEEDISLMDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150

Qy 116 TMRGEIKNGSFNTTISRDEVOKEVALFYKLDVVPIDNN-NTSYRLISCDTSVITQACP 174

Db 151 MVGGEKMGKSPKITTNRGVQKEVALFYELDIPIIDNNSNNRYRLISCDTSVITQACP 210

Qy 175 KISFEPIPIHYCAPAGAILKCKDKKFKNGKPCSNVSTVQCTHGRIPVSTQLLNLSLA 234

Db 211 KISFEPIPIHYCAPAGAILKCKDKKFKNGKPCSNVSTVQCTHGRIPVSTQLLNLSLA 270

Qy 235 EEEVVISDNFTNNAKTIIVOLKESVEINCRPNNNTRKSHIGGRAPFTTGEIIGDIR 294

Db 271 EEEVVISDNFTNNAKTIIVOLKESVEINCRPNNNTRKSHIGGRAPFTTGEIIGDIR 330

Qy 295 QAHGNISRAKNDTLKQIVIKLRQENKTIIVFNHSSGDPDEIVMHSFNCGEFPYCNS 354

AC 078225 PRELIMINARY; PRT; 854 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Env protein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Ba-1;

RA MEDLINE=91289160; PubMed=1905842;

RA Hwang S.S., Boyle T.J., Lyerly H.K., Cullen B.R.;

RT "Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1."

RL Science 253:71-74(1991);

DR EMBL; M63929; AAA75116.1; -

DR InterPro; IPR000328; Env\_GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

SO SEQUENCE 854 AA; 97186 MW; 005FC881B5CBDFD8 CRC64;

Db 331 QAHGNISRAKNDTLKQIVIKLRQENKTIIVFNHSSGDPDEIVMHSFNCGEFPYCNS 390

Qy 355 QLPSTNNNTNTEGNTTITLPCRIKQIIMMVGKAMAPPIRGQIRCSNITG 414

Db 391 QLPSTNNNTNTEGNTTITLPCRIKQIIMMVGKAMAPPIRGQIRCSNITG 450

Qy 415 LLTRDGINENGTEIRPRGGMDRMRSEFYKVKVKEIPLGVAATKCKRRVQREK 474

Db 451 LLTRDGPPEANKTEVFRPGGDMRDNRSELYKVKVKEIPLGVAATKCKRRVQREK 510

Qy 475 VGIGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOQRLQTLTWG 534

Db 511 VGIGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOQRLQTLTWG 570

Qy 535 IKQIQAARVLAVERLYGDOQLLGIMGCGSKLICCTAVPWNASWSKSLDRIMNNMTWM 594

Db 571 IKQIQAARVLAVERLYGDOQLLGIMGCGSKLICCTAVPWNASWSKSLDRIMNNMTWM 630

Qy 595 REIDNYTSEIYTLIEESONOEKNEQELBELDKKASIMNPDITNMLMY 643

Db 631 REIDNYTSEIYTLIEESONOEKNEQELBELDKKASIMNPDITNMLMY 679

Query Match 88.7%; Score 3069.5; DB 15; Length 855;  
Best Local Similarity 88.1%; Pred. No. 1e-241;  
Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

Qy 2 EKLWTVYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61

Db 31 EKLWTVYYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90

Qy 62 HFNMMKNMVEOMOEEDISLMDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115

Db 91 NFNMMKNMVEOMOEEDISLMDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150

Qy 116 TMRGEIKNGSFNTTISRDEVOKEVALFYKLDVVPIDNN-NTSYRLISCDTSVITQACP 174

Db 151 MVGGEKMGKSPKITTNRGVQKEVALFYELDIPIIDNNSNNRYRLISCDTSVITQACP 210

Qy 175 KISFEPIPIHYCAPAGAILKCKDKKFKNGKPCSNVSTVQCTHGRIPVSTQLLNLSLA 234

Db 211 KISFEPIPIHYCAPAGAILKCKDKKFKNGKPCSNVSTVQCTHGRIPVSTQLLNLSLA 270

Qy 235 EEEVVISDNFTNNAKTIIVOLKESVEINCRPNNNTRKSHIGGRAPFTTGEIIGDIR 294

Db 271 EEEVVISDNFTNNAKTIIVOLKESVEINCRPNNNTRKSHIGGRAPFTTGEIIGDIR 330

Qy 295 QAHGNISRAKNDTLKQIVIKLRQENKTIIVFNHSSGDPDEIVMHSFNCGEFPYCNS 354

AC 003805 PRELIMINARY; PRT; 855 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Env protein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,

RA Beaver B.;

RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.

DR EMBL; M68893; AAA44191.1; -

DR InterPro; IPR000328; Env\_GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

SO SEQUENCE 855 AA; 96988 MW; EC631A3B84180C8D CRC64;



```

Db      331 QAHCHLSRAKMDTINKVITKIREOFNGNTIVFKHSSGGDPRIYVHSFPGCGEPFYCNST 390
Qy      355 QLFNSTWNNTEGSSNNTGNTITLPCRIRKQIINMMQEVGKAMVAPPIRGQIRCSNITGL 414
Db      391 QLFNSTWNNTEGSSNNTGNTITLPCRIRKQIINMMQEVGKAMVAPPIRGQIRCSNITGL 450
Qy      415 LITRDGGINENGTEIFRPGGDMRDNRSEFYKRVKIEPLGVAPTKCRVRVOREKRA 474
Db      451 LITRDGGEDEDNKTEVFRPGGDMRDNRSELYKRVKIEPLGVAPTKCRVRVOREKRA 510
Qy      475 VGIGVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQOHLQTLTWG 534
Db      511 VGIGVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQOHLQTLTWG 570
Qy      535 IKQLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAWSNKSILDRINNMNTWME 594
Db      571 IKQLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAWSNKSILDRINNMNTWME 630
Qy      595 REIDNYTSEIYTLIEESONQOEKNEQELLELDKMASLNMWFDITNMLWY 643
Db      631 REINNYTSIYSLIESONQOEKNEQELLELDKMASLNMWFDITNMLWY 679

```

RESULT 7

```

ID      003811 PRELIMINARY; PRT; 853 AA.
AC      003811
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Surface envelope glycoprotein.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN      NCBI_Taxid=11676;
RP      SEQUENCE FROM N.A.
RC      STRAIN=ADA;
RX      MEDLINE=91195299; PubMed=2014229;
RT      Westervelt P., Gendelman H.E., Ratner L.;
RT      "Identification of a determinant within the human immunodeficiency
RT      virus 1 surface envelope glycoprotein critical for productive
RT      infection of primary monocytes."
RL      Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101(1991).
DR      EMBL; M60472; AAA5065.1; -.
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      Envelope protein.
SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

```

Query Match 88.6%; Score 3068; DB 15; Length 853;  
 Best Local Similarity 88.8%; Pred. No. 1.3e-241;  
 Matches 578; Conservative 25; Mismatches 38; Indels 10; Gaps 6;

```

Qy      1 VEKLVNTVYGVGVKWEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENT 60
Db      30 VENTLVNTVYGVGVKWEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENT 89
Qy      61 EHFNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTMER 119
Db      90 ENFMNKNMVEQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDLNATNTNINSSSGM--R 147
Qy      120 GEIKKCSNITTSIRDEVOKEAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACKISFE 179
Db      148 GEIKKCSNITTSIRDKVKQYALFYRLDVVPIDNDNTSYRLINCNSTSIITQACKISFE 207
Qy      180 PIPHYCAPAGFAILKCKNDKTFNGKGPCKNVSTVQCTGIRPVVSTOLLNGLAEEVY 239
Db      208 PIPHYCTPAGFAILKCKDKKTFNGKGPCKNVSTVQCTGIRPVVSTOLLNGLAEEVY 267
Qy      240 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGEIIGDIRQAHCN 299

```

```

Db      268 IRSSNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGEIIGDIRQAHCN 327
Qy      300 ISRAKMNLTQKQIVTKLEOF-ENKTIYFNHSSGGDPRIVHSPFCGEPFYCNSTOLF 358
Db      328 ISRTKMNLTQKQIVTKLEOFNGNTIVFNHSSGGDPRIVHSPFCGEPFYCNSTOLF 387
Qy      359 STWNN---NTEGSSNTEGN-TITLPCRIRKQIINMMQEVGKAMVAPPIRGQIRCSNITG 413
Db      388 STWNNNGWNLQSGNTEGNITITLPCRIRKQIINMMQEVGKAMVAPPIRGQIRCSNITG 447
Qy      414 LITRDGGINENGTEIFRPGGDMRDNRSEFYKRVKIEPLGVAPTKCRVRVOREKRA 473
Db      448 LITRDGGSNTSGSSEIFRPGGDMRDNRSELYKRVKIEPLGVAPTKCRVRVOREKRA 507
Qy      474 AVG-IGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQOHLQTL 532
Db      508 AVGTIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQOHLQTL 567
Qy      533 WGIQLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAWSNKSILDRINNMNTWME 592
Db      568 WGIQLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAWSNKSILDRINNMNTWME 627
Qy      593 WEREIDNYTSEIYTLIEESONQOEKNEQELLELDKMASLNMWFDITNMLWY 643
Db      628 WEREIDNYTGLIYTLIEESONQOEKNEQELLELDKMASLNMWFDISNMLWY 678

```

RESULT 8

```

ID      09PXW7 PRELIMINARY; PRT; 846 AA.
AC      09PXW7
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22; Last annotation update)
DE      GP160, envelope glycoprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN      NCBI_Taxid=11676;
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93189881; PubMed=8446773;
RT      Schmidmayrrova H., Gayet O., Guettari N., Bolmont C., Hirsch I.,
RT      "Characterization of HIV1-PAR, a macrophage-tropic strain: cell
RT      tropism, virus/cell entry and nucleotide sequence of the envelope
RT      glycoprotein."
RL      Res. Virol. 144:21-26(1993).
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

```

Query Match 88.5%; Score 3063.5; DB 15; Length 846;  
 Best Local Similarity 89.3%; Pred. No. 3.1e-241;  
 Matches 575; Conservative 29; Mismatches 35; Indels 5; Gaps 5;

```

Qy      2 EKLWNTVYGVGVKWEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENT 61.
Db      31 DKLWNTVYGVGVKWEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVILENT 90
Qy      62 EHFNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTMER 120
Db      91 EFMNKNMVEQOMEDIISLWDQSLKPCVKLTPLCVTLNCKDLNATNTNINSSSGM--R 150
Qy      121 EIKKCSNITTSIRDEVOKEAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACKISFE 180
Db      151 EIKKCSNITTSIRDKVKQYALFYRLDVVPIDNDNTSYRLINCNSTSIITQACKISFE 210
Qy      181 IPIHYCAPAGFAILKCKNDKTFNGKGPCKNVSTVQCTGIRPVVSTOLLNGLAEEVY 240

```

```

Db 211 IPIHYCAPAGFAILKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 270
Qy 241 RSDNFTNNAKTIIVQLKESVINCITRPNNTNRKSIHIGPGAFTTGGIIGDIOAHONI 300
Db 271 RSDNFTNNAKTIIVQLNESVEIKCTRPNNNTNRKSIPIGPGAFYTTGIIIDIROAHCTI 330
Qy 301 SRKANDTLKOIVIKLRQFENKTIIVFNHSGGDEIYVHSGEGEFYCNSTOLFNST 360
Db 331 SKTMENTFKOIVKLRQYKNTIIVFQSSGDEIYTHFPNCGGEFFYCNSTOLFNST 390
Qy 361 MNNTGSSNNTGNTITLPCRIRKOIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 420
Db 391 W-NDTESSNDTE-RTITLPCRIRKOIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 448
Qy 421 G-INENGTIFRPGGDMRDNMRSEFYKVKYKIEPLGVAFTKCRRVVQREKAVGIGA 479
Db 449 GNNNNNGTEIFRPGGDMRDNMRSELYKVKYKIEPLGVAFTKARRVQREKAVGIGA 508
Qy 480 VFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLRAIEAOQRLQLTWGIKQLO 539
Db 509 MFLGFLGAAGSTMGARRTTLTVQARLLLSGIVQOQNNLRAIEAOHLLQLTWGIKQLO 568
Qy 540 ARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWEEREIDN 599
Db 569 ARVLAVERYLKQQLIGWCSGSKICTTAVPWNASWSNKSID-INNMTWMEWEEREIDN 627
Qy 600 YTSIYTLIEESONQOEKNEOELLELDKWSLMMNFDTINMLMY 643
Db 628 YTNLIYTLIEESONQOEKNEDELLELDKWSLMMNFDTIRMLMY 671

```

## RESULT 9

```

ID 003806 PRELIMINARY; PRT; 838 AA.
AC 003806;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
   Submitted (JUN-1991) to the EMBL/Genbank/DBJ databases.
DR EMBL; M68894; AAA44196.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
SQ SEQUENCE 838 AA; 95329 MW; 11044B1CA10CFB4D CRC64;

```

Query Match 88.4%; Score 3059.5; DB 15; Length 838;

Best Local Similarity 88.8%; Pred. No. 6.4e-241; Indels 9; Gaps 1;

```

Matches 570; Conservative 25; Mismatches 38; Indels 9; Gaps 1;
Qy 2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 61
Db 31 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 90
Qy 62 HFNWKNMNMVEMQEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERGE 121
Db 91 NFNWKNMNMVEMQEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERGE 141
Qy 122 IKNSFNITTSIRUVOKEVALFYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFEPI 181
Db 142 MKNSFNITTSIRUVOKEVALFYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFEPI 201
Qy 182 PIHYCAPAGFAILKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 241

```

```

Db 202 PIHYCAPAGFAILKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 261
Qy 242 SDNFTNNAKTIIVQLKESVINCITRPNNTNRKSIHIGPGAFTTGGIIGDIOAHONI 301
Db 262 SENFTNNAKTIIVQLNESVEIKCTRPNNNTNRKSIPIGPGAFTTGGIIGDIOAHONI 321
Qy 302 RAKANDTLKOIVIKLRQFENKTIIVFNHSGGDEIYVHSGEGEFYCNSTOLFNST 361
Db 322 RAKANDTLKTIIVIKLRQFENKTIIVFQSSGDEIYTHFPNCGGEFFYCNSTOLFNST 381
Qy 362 MNNTGSSNNTGNTITLPCRIRKOIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 421
Db 382 NVTESNNTVANNITTLPCRIRKOIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 441
Qy 422 INENGTIFRPGGDMRDNMRSEFYKVKYKIEPLGVAFTKCRRVVQREKAVGIGA 481
Db 442 PEDKTEVFRPGGDMRDNMRSELYKVKYKIEPLGVAFTKARRVQREKAVGIGA 501
Qy 482 LGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLRAIEAOQRLQLTWGIKQLOAR 541
Db 502 LGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLRAIEAOHLLQLTWGIKQLOAR 561
Qy 542 VLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWEEREIDNT 601
Db 562 VLAVERYLKQQLIGWCSGSKICTTAVPWNASWSNKSINKIMDNTWMEWEEREIDNT 621
Qy 602 SEIYTLIEESONQOEKNEOELLELDKWSLMMNFDTINMLMY 643
Db 622 SIYSLIEESONQOEKNEOELLELDKWSLMMNFDTIRMLMY 663

```

## RESULT 10

```

ID 056110 PRELIMINARY; PRT; 851 AA.
AC 056110;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SEWHS.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
   "Diversity of the human immunodeficiency virus type 1 envelope
   glycoprotein in San Francisco Men's Health Study participants.",
   AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025756; AACC40593.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000173; GAP dhodrogenase.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDAC83EE6464531 CRC64;

```

Query Match 88.4%; Score 3059; DB 15; Length 851;

Best Local Similarity 88.9%; Pred. No. 7.2e-241; Indels 8; Gaps 4;

```

Matches 576; Conservative 28; Mismatches 36; Indels 8; Gaps 4;
Qy 2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 61
Db 31 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 90
Qy 62 HFNWKNMNMVEMQEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERG 120

```

```

Db      91  NFNMAKNMVEOMEDIISLDOSLKPCKVLTPLCVTLNCLDRLDNTTSSGGETMERG 150
Qy      121  EIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151  EIKNCSFNITTSIRKVKQKEYALLHLKLDVVPID-NDTSIRLVSCNITSVITQACPKISFEP 209
Qy      181  IPIHICAPAGPAILKCNCKTFNGKGPCKNVSTVQCTHGIRPVSTOLLINGSLAEVEVI 240
Db      210  IPIHICAPAGPAILKCNCKTFNGKGPCKNVSTVQCTHGIRPVSTOLLINGSLAEVEVI 269
Qy      241  RSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGBRAFYTTGELIIGIRQAHCI 300
Db      270  RSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGBRAFYTTGELIIGIRQAHCI 329
Qy      301  SPAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIVMHSFNCGEFFYCSTOLFENST 360
Db      330  SPAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIVMHSFNCGEFFYCSTOLFENST 389
Qy      361  WN-----NNTGSSNTEGNTITLPCRIKQIIMMOEVRGKAMYAPPIRQIRCSNITGL 415
Db      390  WSNSTNMTDTEGVNNT-GENITLPCRIKQIIMMOEVRGKAMYAPPIRQIRCSNITGL 448
Qy      416  LTRDGINENGTETIRPFGGDMRDMRSFYYKYKYEPLGVAAPKCRVVOBKRAV 475
Db      449  LTRDGINENGTETIRPFGGDMRDMRSFYYKYKYEPLGVAAPKCRVVOBKRAV 508
Qy      476  GIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQOONNLRALIEAQORMLQJTWGI 535
Db      509  GIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQOONNLRALIEAQORMLQJTWGI 568
Qy      536  KOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDRIIMNMWMEWR 595
Db      569  KOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDRIIMNMWMEWR 628
Qy      596  EIDNTSEYTLIESONOEKNEOELELDKMASLMMNFDITNMWY 643
Db      629  EIDNTSEYTLIESONOEKNEOELELDKMASLMMNFDITNMWY 676

```

## RESULT 11

```

ID      077989  PRELIMINARY; PRT; 863 AA.
AC      077989;
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE      Envelope glycoprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SF128A;
RX      MEDLINE=91056585; PubMed=2243391;
RA      Liu Z.-Q., Wood C., Levy J.A., Cheng-Mayer C.;
RT      "The viral envelope gene is involved in macrophage tropism of a human
RL      immunodeficiency virus type 1 strain isolated from brain tissue.";
J. Virol. 64:6148-6153(1990).
DR      EMBL; M95292; AAA44331.1; -
DR      InterPro; IPR000328; Env_Gp11.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 863 AA; 98222 MW; 0157397FAB84C5D7 CRC64;

```

```

Query Match      88.0%; Score 3048; DB 15; Length 863;
Best Local Similarity 87.0%; Pred. No. 5.8e-240;
Matches 574; Conservative 26; Mismatches 40; Indels 20; Gaps 5;

```

```

2 EKLAATVYGVPMKEATTTLCASDAKAYDTEVHNNWATACVPTDPNPOEVLNTE 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      31  EKLAATVYGVPMKEATTTLCASDAKAYDTEVHNNWATACVPTDPNPOEVLNTE 90
Qy      62  HFNMAKNMVEOMEDIISLDOSLKPCKVLTPLCVTLNCKD-----VNAATNTNDSGTM 117
Db      91  NFNMAKNMVEOMEDIISLDOSLKPCKVLTPLCVTLNCLDRLDNTTSSGGETMERG 149
Qy      118  -----ERGEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQ 172
Db      150  SLSRGETGEIKNCSFNITTSIRKVKQKEYALFYKLDVVPIDNNNTSYRLINCNSTITQ 209
Qy      173  CPKISFEPPIHYCAPAGPAILKCNCKTFNGKGPCKNVSTVQCTHGIRPVSTOLLINGSL 232
Db      210  CPKISFEPPIHYCAPAGPAILKCNCKTFNGKGPCKNVSTVQCTHGIRPVSTOLLINGSL 269
Qy      223  LAEEVVRSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGBRAFYTTGELIIGIRQAH 292
Db      270  LAEEVVRSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGBRAFYTTGELIIGIRQAH 329
Qy      293  IROAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIVMHSFNCGEFFYC 352
Db      330  IROAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIVMHSFNCGEFFYC 389
Qy      353  STOLFENSTWN-----NNTGSSNTEGNTITLPCRIKQIIMMOEVRGKAMYAPPIRG 403
Db      390  STOLFENSTWN-----NNTGSSNTEGNTITLPCRIKQIIMMOEVRGKAMYAPPIRG 448
Qy      404  QIRCSNITGLLTRDGINENGTETIRPFGGDMRDMRSFYYKYKYEPLGVAAPKCRV 463
Db      449  QIRCSNITGLLTRDGINENGTETIRPFGGDMRDMRSFYYKYKYEPLGVAAPKCRV 508
Qy      464  KRVVOREKRAVIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQOONNLRALIEA 523
Db      509  KRVVOREKRAVIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQOONNLRALIEA 568
Qy      524  QORMLQJTWGIKOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDR 583
Db      569  QORMLQJTWGIKOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDR 628
Qy      584  IMNMWMEWREREIDNTSEYTLIESONOEKNEOELELDKMASLMMNFDITNMWY 643
Db      629  IMNMWMEWREREIDNTSEYTLIESONOEKNEOELELDKMASLMMNFDITNMWY 688

```

## RESULT 12

```

ID      09WJMS  PRELIMINARY; PRT; 868 AA.
AC      09WJMS;
DT      01-NOV-1999 (TREMBlrel. 12, Created)
DT      01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT      01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE      Envelope polyprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Pang G., Weiher B., Chappey C., Visseky A., Townsend L., Wang Q.,
RA      Burger H.;
RT      "Complete Plasma HIV-1 Sequence: Sp1-Promoter Deletion Can lead to
RL      Non-Progressive Infection.";
Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
DR      EMBL; U69584; AAD10875.1; -
DR      InterPro; IPR000328; Env_Gp41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 868 AA; 98569 MW; F2D138F0F8BEFC02 CRC64;

```

```

Query Match      88.0%; Score 3045.5; DB 15; Length 868;
Best Local Similarity 87.7%; Pred. No. 9.4e-240;
Matches 582; Conservative 22; Mismatches 37; Indels 23; Gaps 7;

```

```

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61
D 31 EQLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90
QY 62 HFNWKNWVQWQEDDIISLWDSLKPCVKLTPLCVTLNCKD-----VNAITNTDSE 114
D 91 NFNWKNWVQWQEDDIISLWDSLKPCVKLTPLCVTLNCKDKNATVKAATNTNNSGW 150
QY 115 GTRMERGEIKNCSFNITTSIRDEVOKEVALFYKLDVPIIDNNNTSYRLISCDT 166
D 151 GGMERGEIKNCSFNITTSIRDEVOKEVALFYKLDVPIIDNNNTSYRLISCDT 210
QY 167 SVITQACPKISFEPPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGRIPVSTQ 226
D 211 SVITQACPKISFEPPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGRIPVSTQ 270
QY 227 LLLNGSLAEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPGRAYTT 286
D 271 LLLNGSLAEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPGRAYTT 330
QY 287 GEIIGDIRQAHCTNIRTEMNNTLAKITEKREQFENNTIIVFNHSSGGDEPIVHSPICG 390
D 331 GEIIGDIRQAHCTNIRTEMNNTLAKITEKREQFENNTIIVFNHSSGGDEPIVHSPICG 390
QY 346 GEFYCNSTQLFNSTWN---NNTGSSNTEGNTITLPCRKOIIMWQEVGKAMVAPPIR 402
D 391 GEFYCNSTQLFNSTWNSTGNISSDTERN--ITLPCRKOIIMWQEVGKAMVAPPIR 449
QY 403 GOIRCSSNITGLLTRDGIN--ENGTEIPRPGGDMRDNMRSEFYKVKVIEPLGVAP 460
D 450 GOIRCSSNITGLLTRDGSNTDENRTEIPRPGGDMRDNMRSEFYKVKVIEPLGVAP 509
QY 461 TKCRVRVQREKRVAG--IGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLR 519
D 510 TKARVRVQREKRVAG--IGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLR 569
QY 520 AIEAQORMLQTLVWGIKOLQARVLAVERYLGDOQLLIGWCSGKLICTTAVPWNASMKN 579
D 570 AIEAQORMLQTLVWGIKOLQARVLAVERYLGDOQLLIGWCSGKLICTTAVPWNASMKN 629
QY 580 SLDRIMNNTWMEWEREIDNTSEIYLLIESONOEKNEBELLDKVASIMNFDITN 639
D 630 SLDEIMNTWMEWEREIDNTSEIYLLIESONOEKNEBELLDKVASIMNFDITN 689
QY 640 WLMY 643
D 690 WLMY 693

```

RESULT 13

092762 PRELIMINARY; PRT; 854 AA.

AC 092762;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SFH56.3;

RC MEDLINE=98178716; PubMed=9519894;

RA McCutchan F.E.; Sanders-Buell E.; Salminen M.O.; Carr J.K.;

RA Sheppard W.H.;

RT "diversity of the human immunodeficiency virus type 1 envelope

RT glycoprotein in San Francisco Men's Health Study participants";

RL AIDS Res. Hum. Retroviruses 14:329-337(1998).

DR EMBL; AF025754; AAC40591.1; -

DR InterPro; IPR000328; Env\_GP41.

```

DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62C8B2D CRC64;

Query Match 87.9%; Score 3044; DB 15; Length 854;
Best Local Similarity 88.3%; Pred. No. 1.2e-239;
Matches 573; Conservative 29; Mismatches 39; Indels 8; Gaps 5;

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61
D 31 EQLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 91
QY 62 HFNWKNWVQWQEDDIISLWDSLKPCVKLTPLCVTLNCKD-----VNAITNTDSE 119
D 91 NFNWKNWVQWQEDDIISLWDSLKPCVKLTPLCVTLNCKDKNATVKAATNTNNSGW 151
QY 120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVPIIDNNNTSYRLISCDTAVITQACPKISFE 179
D 152 GEIKNCSFNITTSIRDEVOKEVALFYKLDVPIIDNNNTSYRLISCDTAVITQACPKISFE 211
QY 180 PIPHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGRIPVSTOLLNGSLAEEV 239
D 212 PIPHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGRIPVSTOLLNGSLAEEV 271
QY 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPGRAYTTGEIIGDIRQAH 299
D 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPGRAYTTGEIIGDIRQAH 331
QY 300 ISRAKNDTLQOIVIKLEQFENNTIIVFNHSSGGDEPIVHSPICGGEFYNSTQLFNS 359
D 332 ISRAKNDTLQOIVIKLEQFENNTIIVFNHSSGGDEPIVHSPICGGEFYNSTQLFNS 391
QY 360 TWNN--NTEGSSNTEGNTITLPCRKOIIMWQEVGKAMVAPPIRGOIRCSSNITGLLT 417
D 392 TWNN--NTEGSSNTEGNTITLPCRKOIIMWQEVGKAMVAPPIRGOIRCSSNITGLLT 451
QY 418 RDGGINENGT---EIPRPGGDMRDNMRSEFYKVKVIEPLGVAPFKRVRVQREKRA 474
D 452 RDGGINENGT---EIPRPGGDMRDNMRSEFYKVKVIEPLGVAPFKRVRVQREKRA 510
QY 475 VGIGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLRAIEAQORMLQTLVW 534
D 511 VGIGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLRAIEAQORMLQTLVW 570
QY 535 IKOLQARVLAVERYLGDOQLLIGWCSGKLICTTAVPWNASMKNSLDRIMNNTWME 594
D 571 IKOLQARVLAVERYLGDOQLLIGWCSGKLICTTAVPWNASMKNSLDRIMNNTWME 630
QY 595 REIDNTSEIYLLIESONOEKNEBELLDKVASIMNFDITNLMY 643
D 631 REIDNTSEIYLLIESONOEKNEBELLDKVASIMNFDITNLMY 679

```

RESULT 14

09YPS0 PRELIMINARY; PRT; 860 AA.

AC 09YPS0;

DT -01-MAY-1999 (TREMBLrel. 10, Created)

DT -01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Envelope polypeptide.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC Fung C.; Weiser B.; Chappey C.; Visosky A.; Townsend L.; Wang Q.;

RC Burger H.;

RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to

RT Non-Progressive Infection.";



**THIS PAGE BLANK (US 710)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:24:08 ; Search time 34.4829 Seconds  
(without alignments)  
2665.163 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129  
Sequence: 1 VEKLMVTYYGVVWKKEAT.....ELDKWASIMNPFDTNLMY 579

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A.GeneSeq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3129	100.0	579	22	AAB61506
2	3056	97.7	643	22	AAB61505
3	2923	93.4	625	22	AAB61507
4	2800.5	89.5	855	18	AAW1581
5	2800.5	89.5	855	20	AAW8813
6	2782	88.9	883	22	AAB82761
7	2776.5	88.7	850	16	AAK67724
8	2771	88.6	619	23	AAU75156
9	2771	88.6	646	23	AAU75155
					N-terminal mutant
					Mediated full-length

10	2764	88.3	847	21	AAV97073
11	2759	88.2	842	24	ABU66565
12	2749	87.9	842	23	ABU66211
13	2735.5	87.4	851	9	AAW80967
14	2725	87.1	857	16	AAW67725
15	2721.5	87.0	868	7	AAW60422
16	2719	86.9	856	22	AAW85999
17	2718.5	86.9	863	14	AAW43869
18	2718.5	86.9	868	23	AAO1389
19	2717.5	86.8	868	7	AAW60063
20	2715	86.8	856	7	AAW61514
21	2715	86.8	856	20	AAW89325
22	2714	86.7	855	19	AAW53112
23	2714	86.7	855	21	AAW77298
24	2714	86.7	855	21	AAW77302
25	2714	86.7	855	24	ABU57550
26	2714	86.7	855	24	ABU57553
27	2714	86.7	863	7	AAW61509
28	2714	86.7	863	13	AAW29706
29	2713.5	86.7	855	12	AAW14905
30	2713	86.7	856	7	AAW60131
31	2712	86.7	854	21	AAW10697
32	2712	86.7	854	21	AAW10053
33	2712	86.7	854	22	AAW86199
34	2712	86.7	854	22	AAW67277
35	2712	86.7	854	23	ABG73663
36	2711	86.6	856	13	AAW25940
37	2709	86.6	863	7	AAW60349
38	2708.5	86.6	860	18	AAW31284
39	2707.5	86.5	880	10	AAW23333
40	2707.5	86.5	880	20	AAW73332
41	2707.5	86.5	880	21	AAW14844
42	2706.5	86.5	856	16	AAW67726
43	2705.5	86.5	855	19	AAW43069
44	2704.5	86.4	901	8	AAW70665
45	2702	86.4	856	21	AAV97072

## ALIGNMENTS

RESULT 1	
AAB61506	
ID	AAB61506 standard; Protein; 579 AA.
AC	AAB61506;
DT	05-APR-2001 (first entry)
DE	HIV-1 delcaviv2* SOS gp140 glycoprotein.
KM	gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp141; gp140.
KW	
OS	Human immunodeficiency virus type 1.
PN	W0200100648-A1.
PD	04-JAN-2001.
PF	23-JUN-2000; 2000MO-US17267.
PR	25-JUN-1999; 99US-0340992.
XX	
XX	(PROG-) PROGENICS PHARM INC.
XX	(AAO-) AARON DIAMOND AIDS RES CENT.
XX	Binley JM, Schuelke N, Olson WC, Madden PJ, Moore JP;
XX	WPI; 2001-122993/13.
XX	N-PSDB; AAF28582.
XX	
XX	New viral envelope proteins, useful for producing vaccines to treat

Variant HIV-1 SF16  
Human immunodeficiency virus isolate SF  
HIV-1 protein HT6.  
gp120 from the HIV  
Sequence of LAV V1  
Amino acid sequence  
HIV-1 Env-LOR g  
Lymphadenopathy-as  
HIV-1 env gene  
Sequence of env  
HIV-1 env protein  
ENV protein contai  
HIV-1 (ATCC CRL 85  
HIV-1 (ATCC CRL 85  
AIDS associated re  
Sequence of ARV-2  
env gene decoded f  
HIV-1 BA-L clone-e  
Sequence of the A1  
HIV-1 env protein.  
HIV-1 coat protein  
HIV gp41 DNA SEQ I  
Protein encoded by  
HIV-1 NL4-3 gp120  
Modified HIV env g  
HIV-III virus (HI  
HIV-SF2 virus gp12  
Human immunodeficiency  
HIV envelope prote  
gp120 from the HIV  
HIV-1 gp120 protei  
Sequence encoded b  
Wild type HIV-1 HX

CC The present invention relates to a viral envelope protein. The viral  
CC envelope protein comprises a viral surface protein (e.g. glycoprotein  
CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in  
CC which the viral envelope protein contains one or more amino acid sequence  
CC mutations that enhance the stability of the complex formed between the  
CC viral surface and transmembrane proteins. The viral envelope protein can  
CC be used in the treatment of viral infection e.g. HIV-1 infection. The  
CC present sequence is HIV-1 SOS gp140, which was used in the present  
CC invention.

QY 357 GINENGTEIFRPGGDMRDNRSELYKKVKIEPLGVAPTCKKRRVQREKRAVGIGAV 416



```

Db      421 GINENGTEIFRPGGDMRNMRSSEFYKYVXVIEPIEGLVAPTKCKRRVQREKRAVGIGAV 480
Qy      417 FLGFLGAAGSTMGASMTLTVOARLLLSGIVOOQNNLLRAIEAQRMQLTVMGIKOLQA 476
Db      461 FLGFLGAAGSTMGASMTLTVOARLLLSGIVOOQNNLLRAIEAQRMQLTVMGIKOLQA 540
Qy      477 RVLAVERYLGDQQLLGIGWCSGKLICTTAVPNNASWSNKSILDRIMNNMTMEREIDNY 536
Db      541 RVLAVERYLGDQQLLGIGWCSGKLICTTAVPNNASWSNKSILDRIMNNMTMEREIDNY 600
Qy      537 TSEIYTLIEESONQOEKNEOELELDKMASLNNMFDTNNLMY 579
Db      601 TSEIYTLIEESONQOEKNEOELELDKMASLNNMFDTNNLMY 643

RESULT 3
AAB61507
ID      AAB61507 standard; Protein; 625 AA.
Qy      AAB61507;
Db      AAB61507;
Dt      05-APR-2001 (first entry)
De      HIV-1 deltaV3 SOS gp140 glycoprotein.
Kx      gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp14;
Kw      gp140.
Xx      Human immunodeficiency virus type 1.
Xx      WO200100648-A1.
Xx      PD      04-JAN-2001.
Xx      PF      23-JUN-2000; 2000MO-US17267.
Xx      PR      25-JUN-1999; 99US-0340992.
Xx      PA      (PROG-) PROGENICS PHARM INC.
Xx      PA      (AARO-) AARON DIAMOND AIDS RES CENT.
Xx      PI      Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;
Xx      DR      MPI; 2001-122993/13.
Xx      DR      N-PSDB; AAF28583.
Xx      PR      New viral envelope proteins, useful for producing vaccines to treat
Xx      human immunodeficiency virus-1 infections, comprises amino acid
Xx      complex is more stable -
Xx      PS      Disclosure; Fig 15; 109pp; English.
Xx      CC      The present invention relates to a viral envelope protein. The viral
Xx      CC      envelope protein comprises a viral surface protein (e.g. glycoprotein
Xx      CC      gp120) and a corresponding viral transmembrane protein (e.g. gp41) in
Xx      CC      which the viral envelope protein contains one or more amino acid sequence
Xx      CC      mutations that enhance the stability of the complex formed between the
Xx      CC      viral surface and transmembrane proteins. The viral envelope protein can
Xx      CC      be used in the treatment of viral infection e.g. HIV-1 infection. The
Xx      CC      present invention is HIV-1 deltaV3 SOS gp140, which was used in the
Xx      CC      present invention.
Xx      SQ      Sequence 625 AA;
Qy      Query Match 93.4%; Score 2923; DB 22; Length 625;
Qy      Best Local Similarity 86.0%; Pred. No. 8.2e-136;
Qy      Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;
Db      1 VEKLMVTVYVGVVWKEATITTLFCASDAKAYDEVVHNVATHACVPTDNPQEVLENT 60
Db      1 VEKLMVTVYVGVVWKEATITTLFCASDAKAYDEVVHNVATHACVPTDNPQEVLENT 60

```

```

Qy      61 EHFNMKNMVEQOMEDIIISLMDOSLKPCVKLTPLCA----- 98
Db      61 EHFNMKNMVEQOMEDIIISLMDOSLKPCVKLTPLCA----- 98
Qy      99 -----GCPTSVITQACPKISEP 116
Db      121 EIKNCSFNITTSIRDEVOKEYALFYKLDVVAIDNNNTSYRLISCOTSVITQACPKISEP 180
Qy      117 IPIHCAAGAPAILKCNKRTNGKPCGVSTVQCTHGRIPVVSQLLNGSLAEEVYI 176
Db      181 IPIHCAAGAPAILKCNKRTNGKPCGVSTVQCTHGRIPVVSQLLNGSLAEEVYI 240
Qy      177 RSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHNI 236
Db      241 RSDNFTNAKTIIVOLKESVEINCTRPNNN-----AGDIRQAHNI 282
Qy      237 SRAKANDTLKQIVIKLRQEPENKTIIVFNHSSGGDEIVMHSFNCGSEFFYCNSLTQFNST 296
Db      283 SRAKANDTLKQIVIKLRQEPENKTIIVFNHSSGGDEIVMHSFNCGSEFFYCNSLTQFNST 342
Qy      297 MNNTGSGNNTGNTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCSSNTGLLTRDG 356
Db      343 MNNTGSGNNTGNTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCSSNTGLLTRDG 402
Qy      357 GINENGTEIFRPGGDMRNMRSSEFYKYVXVIEPIEGLVAPTKCKRRVQREKRAVGIGAV 416
Db      403 GINENGTEIFRPGGDMRNMRSSEFYKYVXVIEPIEGLVAPTKCKRRVQREKRAVGIGAV 462
Qy      417 FLGFLGAAGSTMGASMTLTVOARLLLSGIVOOQNNLLRAIEAQRMQLTVMGIKOLQA 476
Db      463 FLGFLGAAGSTMGASMTLTVOARLLLSGIVOOQNNLLRAIEAQRMQLTVMGIKOLQA 522
Qy      477 RVLAVERYLGDQQLLGIGWCSGKLICTTAVPNNASWSNKSILDRIMNNMTMEREIDNY 536
Db      523 RVLAVERYLGDQQLLGIGWCSGKLICTTAVPNNASWSNKSILDRIMNNMTMEREIDNY 582
Qy      537 TSEIYTLIEESONQOEKNEOELELDKMASLNNMFDTNNLMY 579
Db      583 TSEIYTLIEESONQOEKNEOELELDKMASLNNMFDTNNLMY 625

RESULT 4
AAM11581
ID      AAM11581 standard; Protein; 855 AA.
Qy      AAM11581;
Db      AAM11581;
Dt      25-MAR-2003 (updated)
Dt      25-MAR-1997 (first entry)
Xx      DE      Human Immunodeficiency Virus-1 strain BA-L envelope protein.
Xx      KW      Acquired immune deficiency syndrome; AIDS; envelope protein;
Xx      KW      env gene; vaccine.
Xx      OS      Human immunodeficiency virus type 1 (strain BA-L).
Xx      FN      US5576000-A.
Xx      PD      19-NOV-1996.
Xx      PF      15-FEB-1995; 95US-0388809.
Xx      PR      17-OCT-1990; 90US-0599491.
Xx      PR      25-FEB-1993; 93US-0022835.
Xx      PR      15-FEB-1995; 95US-0388809.
Xx      PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
Xx      PI      Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
Xx      PI      Popovic M, Reitz MS;
Xx      XX

```

DR WPI; 1997-011206/01.  
 DR N-PSDB; AAT58551.  
 PT New isolated envelope protein of HIV-1 strain BA-L and recombinant  
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,  
 PT typical of US clinical isolates  
 PS  
 XX Claim 1; Fig 9; 86pp; English.  
 XX  
 CC A HindIII fragment of unintegrated viral DNA representing the BA-L  
 CC genome was cloned into lambda phage Charon 28 DNA from total DNA of  
 CC peripheral blood lymphocytes infected with and producing HIV-1(BA-L).  
 CC A positive clone was selected by hybridisation using a HIV-1 env  
 CC probe. This clone, designated BA-L1, was found to contain the  
 CC entire gene for the envelope protein on a 2.8 kb HindIII-XbaI  
 CC fragment and a 0.4 kb EcoRI-HindIII fragment. When cloned together  
 CC these fragments comprise the env gene, as well as the coding regions  
 CC for rev and the rev-responsive element of env, both necessary for  
 CC efficient expression in eukaryotic cells. The claimed recombinantly  
 CC produced envelope protein can be used as an immunogen for raising  
 CC antibodies against HIV.  
 CC (updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 855 AA;  
 Query Match 89.5%; Score 2800.5; DB 18; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 1e-129;  
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;  
 QY 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 61  
 DB 31 EKLWTVYVYGVVWKEATTTLFCASDRKAYDTEVHNWATHACVPTDPNQEVLKNVTE 90  
 QY 62 HFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC-----CPTSVTQACP 110  
 DB 91 NFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC-----CPTSVTQACP 150  
 QY 97 --GAG-----CPTSVTQACP 110  
 DB 151 MVGGEMKNCSEFNTTNRGKVKQYALFYKLDIAPIDNNSNNRRLISCTSVITQACP 210  
 QY 111 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLA 170  
 DB 211 KVSFEPIPIHYCAPAGFAILKCKDKKFKNGKPCSTVSTVQCTHGIRPVVSTQLLNGSLA 270  
 QY 171 EEEVIRSDNFTNNAKTIIVQAKSEVEINCRPNNTKSHIGGRAPYTTGELIGDIR 230  
 DB 271 EEEVIRSANPADNAKVIIVQAKSEVEINCRPNNTKSHIGGRAPYTTGELIGDIR 330  
 QY 231 QAHGNIISPAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIIVHSGGGEFFYCNST 290  
 DB 331 QAHGNIISPAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIIVHSGGGEFFYCNST 390  
 QY 291 QLFNSTMNNTSGSNTEGNTTTLPCRIKQIIMNQEVKAMYPADIRGQIRCSNITGL 350  
 DB 391 QLFNSTMNNTSGSNTEGNTTTLPCRIKQIIMNQEVKAMYPADIRGQIRCSNITGL 450  
 QY 351 LITRDGGINENGTEIFRPGGGMRRNMSSELYKYRVVLTIEPLGVAFTCKRRVQREGRA 410  
 DB 451 LITRDGGINENGTEIFRPGGGMRRNMSSELYKYRVVLTIEPLGVAFTCKRRVQREGRA 510  
 QY 411 VGIGAVFIFGLGAAGSTWGAASMTLTVQARLLLSGIYVOOONNLALAIQAQOQMLQTVWG 470  
 DB 511 VGIGAVFIFGLGAAGSTWGAASMTLTVQARLLLSGIYVOOONNLALAIQAQOQMLQTVWG 570  
 QY 471 IKQIQARVLAVERLYGDOOLGIGWCSGKLICTTAVPNNASGNSKSLDRINNNMTWME 530  
 DB 571 IKQIQARVLAVERLYGDOOLGIGWCSGKLICTTAVPNNASGNSKSLDRINNNMTWME 630  
 QY 531 REINDYTSBITTLIESONQOEKNEOELLEDDKMASLWNPEDITWMLWY 579  
 DB 631 REINDYTSBITTLIESONQOEKNEOELLEDDKMASLWNPEDITWMLWY 679

RESULT 5  
 ID AAM88113  
 XX AAM88113 standard; Protein: 855 AA.  
 AC AAM88113;  
 DT 09-APR-1999 (first entry)  
 XX  
 DE Env protein of the BA-L strain of Human immunodeficiency virus type 1.  
 XX  
 XX HIV-1; HIV-1 strain BA-L; env protein; vaccine;  
 KW immunotherapy; HIV infection; immunogen; HIV-1 diagnosis.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 FH Key  
 FT Misc-difference 11  
 FT Location/Qualifiers  
 FT  
 XX US5869313-A.  
 XX  
 PD 09-FEB-1999.  
 XX  
 PF 14-MAY-1996; 96US-0647714.  
 XX  
 PR 17-OCT-1990; 90US-0599491.  
 PR 25-FEB-1993; 93US-0022835.  
 PR 15-FEB-1995; 95US-0388809.  
 PR 14-MAY-1996; 96US-0647714.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;  
 PI Popovic W, Reitz MS;  
 XX  
 DR WPI; 1999-152779/13.  
 DR N-PSDB; AAX04767.  
 XX  
 PT DNA encoding env protein of the human immune deficiency virus  
 PT isolate BA-L - useful for producing protein for use in vaccines, as  
 PT assay reagent and to generate antibodies  
 XX  
 PS Example 1; Fig 9A-C; 87pp; English.  
 XX  
 CC The present sequence represents the envelope protein of the BA-L  
 CC (ATCC 40890) strain of Human immunodeficiency virus type 1 (HIV-1)  
 CC strain MN-ST1. BA-L is more typical of United States isolates of  
 CC HIV-1 than previously known strains. Recombinant, complete env protein  
 CC of the BA-L strain is used as a vaccine component and for immunotherapy  
 CC of existing HIV infections, to detect HIV-specific antibodies, e.g. in  
 CC donated blood, and as an immunogen to raise specific antibodies, for  
 CC HIV-1 diagnosis.  
 CC  
 XX  
 SQ Sequence 855 AA;  
 Query Match 89.5%; Score 2800.5; DB 20; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 1e-129;  
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;  
 QY 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 61  
 DB 31 EKLWTVYVYGVVWKEATTTLFCASDRKAYDTEVHNWATHACVPTDPNQEVLKNVTE 90  
 QY 62 HFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC-----CPTSVTQACP 110  
 DB 91 NFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC-----CPTSVTQACP 150  
 QY 97 --GAG-----CPTSVTQACP 110  
 DB 151 MVGGEMKNCSEFNTTNRGKVKQYALFYKLDIAPIDNNSNNRRLISCTSVITQACP 210  
 QY 111 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLA 170

```

Db      211 KSFSEPIPIHYCAPAGFALKCKDKKFNKGPCSTNVSTVQCTHGIRPVSTOLLNGSLA 270
Qy      171 EEEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTRSRHIGPRARVYTGELIGDIR 230
Db      271 EEEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTRSRHIGPRARVYTGELIGDIR 330
Qy      231 QAHCHNISRAKAMDLLKQIVIKLREOPENNTIVFNHSSGGDEIYVMSFNGCGEPFYCNST 290
Db      331 QAHCHNISRAKAMDLLKQIVIKLREOPENNTIVFNHSSGGDEIYVMSFNGCGEPFYCNST 390
Qy      291 QLFNSTWNNNTGSNNTEBNTITLPCRKOIINMGEVRAKAYAPRIGQIRCSNITGL 350
Db      391 QLFNSTWNNNTGSNNTEBNTITLPCRKOIINMGEVRAKAYAPRIGQIRCSNITGL 450
Qy      351 LITRDGGINENGTETLPRPGGDMRDWMSRLKYKVKVKEIPGVATPKRRRVQEKKA 410
Db      451 LITRDGGINENGTETLPRPGGDMRDWMSRLKYKVKVKEIPGVATPKRRRVQEKKA 510
Qy      411 VGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQHMLQLTWVG 470
Db      511 VGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQHMLQLTWVG 570
Qy      471 IKQLOARVLAVERYLADQQLIGWCGSKLICCTAVPNNASNSNSLDRIMNNMTWMEWE 530
Db      571 IKQLOARVLAVERYLADQQLIGWCGSKLICCTAVPNNASNSNSLDRIMNNMTWMEWE 630
Qy      531 REIDNYSIETYLIESNOQOEKNEBELLEDKMSLSLMMFDTITWLMY 579
Db      631 REIDNYSIETYLIESNOQOEKNEBELLEDKMSLSLMMFDTITWLMY 679

```

## RESULT 6

AAB82761 standard, Protein; 883 AA.

```

ID      AAB82761 standard, Protein; 883 AA.
AC      AAB82761;
DT      29-OCT-2001. (first entry)
DE      Ancestral HIV-1 group M, subtype B gp160 protein.
KM      HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.
OS      Human immunodeficiency virus type 1.
PN      WO200160838-A2.
PF      23-AUG-2001.
PR      16-FEB-2001; 2001WO-US05288.
PR      18-FEB-2000; 2000US-0183659.
XX      (UNITM ) UNIV WASHINGTON.
PI      Mulline JT, Rodrigo AG, Learn GH, Li F;
DR      WPI: 2001-536565/59.
DR      N-PSDB; AAH26468.
PT      Preparing an ancestral viral amino acid sequence useful as a vaccine
PT      comprises determining a recent ancestor of a circulating virus by
PT      maximum likelihood phylogeny analysis -
XX      Claim 8; Page 54; 89pp; English.
CC      The present sequence is that of an ancestral HIV-1 group M, subtype
CC      B gp160 (env gene product) sequence. The invention provides
CC      compositions and methods for determining ancestral viral gene
CC      sequences and ancestral viral protein sequences for highly diverse
CC      viruses, such as HIV-1. The methods use samples of circulating
CC      viruses to determine an ancestral viral sequence by maximum

```

```

CC      likelihood phylogeny analysis. In the present case, the ancestral
CC      HIV-1 subtype B env sequence (see AAH26468) was determined using 38
CC      B sequences (obtained from 9 different countries) and 3
CC      subtype D (outgroup) sequences. The distances between this ancestral
CC      viral sequence and circulating strains used to determine it were on
CC      average 12.3% (range: 8.0-21.0%) while the available specimens were
CC      17.3% different from each other (range: 13.3-23.2%). Thus, the
CC      ancestor sequence was, on average, more closely related to any given
CC      circulating virus than to any other variant. The ancestral gp160
CC      sequence included a wide variety of immunogenically active peptides
CC      when processed for antigen presentation; nearly all known subtype B
CC      CTL epitope consensus amino acids were represented. Thus, an
CC      immunogenic composition to this subtype B ancestor protein will
CC      elicit broad neutralising antibody against HIV-1 isolates of the
CC      same subtype, and will also elicit a broad cellular response
CC      mediated by antigen-specific T-cells. A claimed vaccine composition
CC      comprises a viral ancestor protein or its immunogenic fragment,
CC      especially one derived from the HIV-1 group M subtype B gp160
CC      ancestral protein.
SQ      Sequence      883 AA:
Query Match      88.9%; Score 2782; DB 22; Length 883;
Best Local Similarity 79.2%; Pred. No. 8.4e-129;
Matches 537; Conservative 14; Mismatches 27; Indels 100; Gaps 7;
Qy      2 EKLMTVYVGVPMVEATTTLPASDARAYDEVNNAWATHCVPTDPPOVULENTE 61
Db      31 EKLMTVYVGVPMVEATTTLPASDARAYDEVNNAWATHCVPTDPPOVULENTE 90
Qy      62 HFNMAKNNMVEQMOEDISLMDQSLKPCVKLTPLC----- 96
Db      91 HFNMAKNNMVEQMHEDIISLMDQSLKPCVKLTPLCLTLCDDTLATNATNTSSATNT 150
Qy      97 ---GAG----- 99
Db      151 TSSGGTMEGEKEIKNCSFNVTTSIRDMQKEVALFYKLDVVPIDNDNNNTNNTSYRL 210
Qy      100 ---CDTSVITQACPKSPFPIPIHYCAPAGFALKCKDRTFNKGCKXNVSTVQCHGRP 157
Db      211 INCNTSVITQACPKSPFPIPIHYCAPAGFALKCKDRTFNKGCKXNVSTVQCHGRP 270
Qy      158 VSTOLLNGSLAESEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTRSRHIGPR 217
Db      271 VSTOLLNGSLAESEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTRSRHIGPR 330
Qy      218 AFTYTGELIGDIRQAHCHNISRAKAMDLLKQIVIKLREOPENK--TIVFNHSSGGDEIYV 275
Db      331 ALYATGKTIIGDIRQAHCHNISRAKAMDLLKQIVIKLREOPENKTIIVFNHSSGGDEIYV 390
Qy      276 HSFNCGGEFFYCNSTQLFNSTW-----NNTBESNNT--EGNTITLPCRKOIINMW 326
Db      391 HSFNCGGEFFYCNSTQLFNSTWTFHFNSTWNNNTESNNAADNDITLPCRKOIINMW 450
Qy      327 EYKAMVAPRIGQIRCSNITGLLTRDGINEN-----GTETLPCRKOIINMW 382
Db      451 EYKAMVAPRIGQIRCSNITGLLTRDGINENNTNTDTEIFRPGGDMRDWMSRLKY 510
Qy      383 KYKVKIPELGVAPTKRRRVQEKRAVG--IGAVFLGFLGAAGSTMGASMTLTVQARL 441
Db      511 KYKVKIPELGVAPTKRRRVQEKRAVGMLGAVFLGFLGAAGSTMGASMTLTVQARL 570
Qy      442 LLSGIVQOQNNLLRAIEAQHMLQLTWVGIIKQLOARVLAVERYLADQQLIGWCGSKLI 501
Db      571 LLSGIVQOQNNLLRAIEAQHMLQLTWVGIIKQLOARVLAVERYLADQQLIGWCGSKLI 630
Qy      CCTAVPNNASNSNSLDRIMNNMTWMEWEIRIDNTYSIETYLIESNOQOEKNEBELLE 561
Db      631 CTTAVPNNASNSNSLDRIMNNMTWMEWEIRIDNTYSIETYLIESNOQOEKNEBELLE 690
Qy      562 DKMASLMMFDTITWLMY 579
Db      691 DKMASLMMFDTITWLMY 708

```

## RESULT 7

AA67724 standard; Protein, 850 AA.

AA67724;

25-MAR-2003 (updated)  
07-SEP-1995 (first entry)

gp120 from the HIV GNE8 isolate.

HIV; human immunodeficiency virus; gp120; glycoprotein;  
GNE clone; GNE8 isolate.

Human immunodeficiency virus type 1.

MO9428929-A1.

22-DEC-1994.

07-JUN-1994; 94WO-US06036.

07-JUN-1993; 93US-0072833.

(GETH ) GENENTECH INC.

Berman PW, Nakamura GR;

WPI; 1995-036112/05.

N-PSDB; AA076018.

Use of HIV gp 120 polypeptide(s) - for developing probes for the  
analysis, prevention and therapy of HIV infection

Claim 23; Page 27-30; 108bp; English.

This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate  
of HIV-1 (human immunodeficiency virus type 1). There are neutralising  
epitopes in the V2 and C4 domains of gp120, in addition to the  
neutralising epitopes in the V3 domain. Although the amino acid sequences  
of the neutralising epitopes in the V2, V4 and C4 domains are variable,  
the amount of variation is highly constrained. This facilitates the  
design of HIV subunit vaccines that can induce antibodies that neutralise  
the most common HIV strains for a given geographic region. This invention  
provides a multivalent gp120 subunit vaccine where the gp120 present in  
the vaccine is from at least two HIV isolates which have different amino  
acid sequences for a neutralising epitope in these regions.  
(Updated on 25-MAR-2003 to correct PN field.)

Sequence . 850 AA;

Query Match 88.7%; Score 2776.5; DB 16; Length 850;  
Best Local Similarity 81.5%; Pred. No. 1.5e-128;  
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

QY 2 EKLMTVYGVVWKEATTTTCASDARVDEYVNNWATHACVETDNPQEVLENTE 61  
DB 31 EKLMTVYGVVWKEATTTTCASDARVDEYVNNWATHACVETDNPQEVLENTE 90  
QY 62 HFNWKNMNVEMOEDTISLMDQSLKPCVKLTPLCGA----- 98  
DB 91 NFNWKNMNVEMOEDTISLMDQSLKPCVKLTPLCGA----- 150  
QY 99 -----GCDTSVITQACPKRISFP 116  
DB 151 EIKNCSFVNTSIRDKMKNVYALFYKLDVPIIDNNTSYRLISCTSVITQACPKRISFP 210  
QY 117 IPIHYCAGAPFALIKCDNTPKSGPCCKNVSVQCTHGIRPVSVQLLNGSLAEVYI 176  
DB 211 IPIHYCAGAPFALIKCDNTPKSGPCCKNVSVQCTHGIRPVSVQLLNGSLAEVYI 270

QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQAHCTI 236  
DB 271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYATGIIIGDIRQAHCTI 330  
QY 237 SRKMNUTLKQIYKLEQFENKTIIVNHSNGDPEIWNHSFNGGFFFCNSQLNST 296  
DB 331 SSTKMNTTLKQIYKLEHFNKTIIVNHSNGDPEIWNHSFNGGFFFCNTPLNST 389  
QY 297 W-----NNTGSSNNTGNTITLPCRIKOIINMGEVGMAYAPPIRGQIRCSNITGL 351  
DB 390 WNTYTYWNTGSSNDT-GRNITLQCRIKOIIINMGEVGMAYAPPIRGQIRCSNITGL 448  
QY 352 LTRDGINNGTEIFRPGGDMRDMRSSELYKVKVIEPLGVAFYCKRVRVQREKRAV 411  
DB 449 LTRDGG-NSSTETFRPGGDMRDMRSSELYKVKVIEPLGVAFYCKRVRVQREKRAV 507  
QY 412 GIGAVFLGFLGAGSTNGAASMTLTVOARLLSGIVQOONNLRAIEAQRMLOLTWGI 471  
DB 508 GIGAVFLGFLGAGSTNGAASMTLTVOARLLSGIVQOONNLRAIEAQRMLOLTWGI 567  
QY 472 KQLQARVLAVERVLDGQOLGIMGSGKLICTTAVPNWASNSKSLDRINNTMTMEMER 531  
DB 568 KQLQARVLAVERVLDGQOLGIMGSGKLICTTAVPNWASNSKSLDRINNTMTMEMER 627  
QY 532 EIDNVTSEIYTLIESQNOQEKNEBELLEDKWASLWNPEDITWLMY 579  
DB 628 EIDNVTSEIYTLIESQNOQEKNEBELLEDKWASLWNPEDITWLMY 675

RESULT 8  
AAU75156  
ID AAU75156 standard; Protein, 619 AA.

AAU75156;

08-MAY-2002 (first entry)

N-terminal mutant of modified HIV-1 SF162DV2 gp140 envelope protein.

Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;  
second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;  
acquired immunodeficiency syndrome; anti-HIV; mutant; mutein; immunogen.  
Human immunodeficiency virus type 1 strain SF162 (clade B).  
Synthetic.

Key Location/Qualifiers  
FT Misc-difference 512  
note= "Encoded by CAGCTC"

MO200200250-A2.

03-JAN-2002.

27-JUN-2001; 2001WO-US20483.

27-JUN-2000; 2000US-214608P.

26-JUN-2001; 2001US-0891609.

(AARO-) AARON DIAMOND AIDS RES CENT.

Stamatatos L, Barnett S, Shrivastava I;

WPI; 2002-110836/17.

N-PSDB; AAS15499.

Immunizing an animal and eliciting an immune response against  
heterologous HIV-1 in an animal, involves administering an immunogen  
comprising modified HIV-1 envelope protein, or DNA or virus encoding  
the protein -  
Claim 5; Fig 17; 62bp; English.

CC The present invention relates to methods for immunising an animal,  
 CC preferably human, against heterologous human immunodeficiency virus  
 CC type 1 (HIV-1), and eliciting a heterologous immune response to HIV-1  
 CC in that animal. The method comprises administering an immunogen having  
 CC at least one modified HIV-1 envelope protein with a deletion in the  
 CC second hypervariable (V2) region or its fragment, or DNA or virus  
 CC encoding the modified HIV-1 envelope protein. The invention discloses  
 CC the polynucleotide and amino acid sequences for the full-length HIV-1  
 CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal  
 CC mutant of SF162DV2. The immunogens of the invention are useful for  
 CC immunising an animal against heterologous HIV-1 strains by eliciting  
 CC neutralising antibodies or protective antibodies in the animal. The  
 CC method can be used to treat acquired immunodeficiency syndrome (AIDS).  
 CC The present sequence represents the N-terminal mutant (deletion of  
 CC amino acids 1-27) of modified HIV-1 SF162DV2 gp140 envelope protein.

SO Sequence 619 AA;

Query Match 88.6%; Score 2771; DB 23; Length 619;

Best Local Similarity 85.6%; Pred. No. 2, 2e-128;

Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

QY 1 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60  
 DB 3 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 62  
 QY 61 EHFNNKNNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 96  
 DB 63 EHFNNKNNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 122  
 QY 97 -----GAG-----CDTSVITQACPKISEPPIPIHYCAPAGAILKCNKDTFNGKGP 142  
 DB 123 GEIKNCSFKVAGKGLNCTSVITQACPKVSEPIPIHYCAPAGAILKCNKDTFNGKGP 182  
 QY 143 CKNVSTVQCTHGIRPVVSTQLLNGLSLAEVVIIRSDNTNNAKTIIVQLKESVEINCTR 202  
 DB 183 CKNVSTVQCTHGIRPVVSTQLLNGLSLAEVVIIRSDNTNNAKTIIVQLKESVEINCTR 242  
 QY 203 PNNNTRKSIHIGPRAFYTTGELIGIRQAHNCNISAKKNDTLKOIVILRQOFENKTI 262  
 DB 243 PNNNTRKSTIGPRAFYTTGELIGIRQAHNCNISAKKNDTLKOIVILRQOFENKTI 302  
 QY 263 FNNHSGGDEIIVHSGFCGSEFFYCNSTQLFNSVNNNNGSNTTNGTITLPCRIKOI 322  
 DB 303 FNNHSGGDEIIVHSGFCGSEFFYCNSTQLFNSVNNNNGSNTTNGTITLPCRIKOI 360  
 QY 323 NNMQVGVKAMVAPPIRGQIRCSNITGLLLTDDGIN-ENGTEIFRPGGDMRDMRSEL 381  
 DB 361 NNMQVGVKAMVAPPIRGQIRCSNITGLLLTDDGIN-ENGTEIFRPGGDMRDMRSEL 420  
 QY 382 YKVKVVKIEPLGVAPTKCRVVOREKRAVIGAVELGFLGAAGSTGMAASMTLVQAAL 441  
 DB 421 YKVKVVKIEPLGVAPTKCRVVOREKRAVIGAVELGFLGAAGSTGMAASMTLVQAAL 480  
 QY 442 LLSGIVQOQNNLLRAIEAOQRMQLTWGIKQLQARVLAVERVLDGQQLIGWCSGKLI 501  
 DB 481 LLSGIVQOQNNLLRAIEAOQHLQLTWGIK-LQARVLAVERVLDGQQLIGWCSGKLI 539  
 QY 502 CCTAVPMAASNSKSLDIRIMNMTWMEWEREIDNTYSSEIYLLIEESQOEKNEDELLEL 561  
 DB 540 CCTAVPMAASNSKSLDIRIMNMTWMEWEREIDNTYLLIEESQOEKNEDELLEL 599  
 QY 562 DKMAASLMMFDTITNLMWY 579  
 DB 600 DKMAASLMMFDTITNLMWY 617

DT 08-MAY-2002 (first entry)  
 XX Modified full-length HIV-1 SF162DV2 gp140 envelope protein.  
 DE  
 XX Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;  
 KW second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;  
 KW acquired immunodeficiency syndrome; anti-HIV; mutant; mutein; immunogen.  
 XX  
 OS Human immunodeficiency virus type 1 strain SF162 (clade B).  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 539  
 FT FT  
 XX W0200200250-A2.  
 PN  
 XX 03-JAN-2002.  
 PD  
 XX 27-JUN-2001; 2001WO-US20483.  
 PF  
 XX 27-JUN-2000; 2000US-214608P.  
 PR 26-JUN-2001; 2001US-0891609.  
 XX (AARO-) AARON DIAMOND AIDS RES CENT.  
 PA Stamatatos L, Barnett S, Shrivastava I;  
 PI WPI; 2002-130836/17.  
 XX N-PSDB; AAS15498.  
 DR  
 XX Immunizing an animal and eliciting an immune response against  
 PT heterologous HIV-1 in an animal, involves administering an immunogen  
 PT comprising modified HIV-1 envelope protein, or DNA or virus encoding  
 PT the protein  
 XX  
 XX Claim 5; Fig 16; 62pp; English.

CC The present invention relates to methods for immunising an animal,  
 CC preferably human, against heterologous human immunodeficiency virus  
 CC type 1 (HIV-1), and eliciting a heterologous immune response to HIV-1  
 CC in that animal. The method comprises administering an immunogen having  
 CC at least one modified HIV-1 envelope protein with a deletion in the  
 CC second hypervariable (V2) region or its fragment, or DNA or virus  
 CC encoding the modified HIV-1 envelope protein. The invention discloses  
 CC the polynucleotide and amino acid sequences for the full-length HIV-1  
 CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal  
 CC mutant of SF162DV2. The immunogens of the invention are useful for  
 CC immunising an animal against heterologous HIV-1 strains by eliciting  
 CC neutralising antibodies or protective antibodies in the animal. The  
 CC method can be used to treat acquired immunodeficiency syndrome (AIDS).  
 CC The present sequence represents the full-length modified HIV-1  
 CC SF162DV2 gp140 envelope protein.

SO Sequence 646 AA;

Query Match 88.6%; Score 2771; DB 23; Length 646;

Best Local Similarity 85.6%; Pred. No. 2, 3e-128;

Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

QY 1 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60  
 DB 3 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89  
 QY 61 EHFNNKNNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 96  
 DB 90 EHFNNKNNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 149  
 QY 97 -----GAG-----CDTSVITQACPKISEPPIPIHYCAPAGAILKCNKDTFNGKGP 142  
 DB 150 GEIKNCSFKVAGKGLNCTSVITQACPKVSEPIPIHYCAPAGAILKCNKDTFNGKGP 209  
 QY 143 CKNVSTVQCTHGIRPVVSTQLLNGLSLAEVVIIRSDNTNNAKTIIVQLKESVEINCTR 202

RESULT 9

AAU75155 ID AAU75155 standard; Protein; 646 AA.

XX AAU75155;

XX

```

Db 210 CTNSTVQCTGIRRVSTQLLNGLSLABEGVIRSENFNTNNAKTIIVQLKESVEINCTR 269
Qy 203 PNNNRKSIHIGPGAFYTTGEIIGDIOAHCHNISRANKNTLKOIVIKLEBQENKTIIV 262
Db 270 PNNNRKSIITIGPGAFYATGDIIGDIOAHCHNISGKMNNTLKOIVTKLOAFENKTIIV 329
Qy 263 FNNSSGDPPEIVMHSFNCGEFFYCNSTOLFNTNNTNNTGEGNTITLPCRKOII 322
Db 330 FKQSSGDPPEIVMHSFNCGEFFYCNSTOLFNTNNTGEGNTITLPCRKOII 387
Qy 323 NMMQEVKAMTAPPIRGQIRGSSNTGLLTRDGIN-ENGTEIFRPGGDMRDNRSEL 381
Db 388 NRMQEVKAMTAPPIRGQIRGSSNTGLLTRDGIN-ENGTEIFRPGGDMRDNRSEL 447
Qy 382 YKRVVVKIEPIGVAPTKRRRVQEKRAVIGAFIPLGAAGSTMAASMTLTVQARL 441
Db 448 YKRVVVKIEPIGVAPTKRRRVQEKRAVIGAFIPLGAAGSTMAASMTLTVQARL 507
Qy 442 LLSGIVQOQNNLLRAIEAQORMLQTVWGIKQLQARVLAVERVLDQQLGIWGGSGKLI 501
Db 508 LLSGIVQOQNNLLRAIEAQORMLQTVWGIKQLQARVLAVERVLDQQLGIWGGSGKLI 566
Qy 502 CTAVPANASWSNKSLEDIRIMNMNTMWEWEREIDNTSEIYTLIESQNOQEKNEQELLE 561
Db 567 CTAVPANASWSNKSLEDIRIMNMNTMWEWEREIDNTSEIYTLIESQNOQEKNEQELLE 626
Qy 562 DKMASLNMWPDITMWMY 579
Db 627 DKMASLNMWPDITMWMY 644

```

## RESULT 10

```

AA97073
ID AA97073 standard; protein; 847 AA.
XX
AC AA97073;
XX
DT 31-OCT-2000 (first entry)
XX
DB Variant HIV-1 SFI62 Env gp160.
XX
KW HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;
KM CD4 binding region; V1/V2 loop; bridging sheet.
OS Human immunodeficiency virus type 1 isolate SFI62.
XX
MO200039303-A2.
XX
06-JUL-2000.
XX
PE 30-DEC-1999; 99WO-US31272.
XX
PR 31-DEC-1998; 98US-0114495.
PR 29-SEP-1999; 99US-0156670.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Hartog K, Martin E;
XX
WP1; 2000-465745/40.
XX
PT Novel modified HIV Env polypeptides useful as immunizing agents and for
PT preparing a vaccine to elicit an immune response against a broad range
PT of HIV subtypes
XX
PS Claim 5; Page 115-117; 139pp; English.
XX
CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope
CC (Env) polypeptides having an amino acid deleted or replaced in the region
CC corresponding to residues 420-436 or 119-123 and 199-210 relative to
CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV
CC strain SFI62, with numbering relative to isolate HXB-2. The Env

```

CC polypeptides are modified so as to expose at least part of the CD4  
 CC binding region. The modified HIV Env polypeptides, coding polynucleotides  
 CC and constructs, further comprising an adjuvant, are used for inducing an  
 CC immune response in an individual. The method involves administering a  
 CC first composition comprising a polynucleotide encoding the Env  
 CC polypeptide in a priming step and administering a second composition  
 CC comprising a modified Env polypeptide as a booster in an amount  
 CC sufficient to induce an immune response in the individual. The first  
 CC and/or second composition further comprises an adjuvant (claimed). The  
 CC intracellularly produced Env polypeptides can be used for a number of  
 CC diagnostic and therapeutic purposes to determine the presence of reactive  
 CC antibodies/and or Env proteins in a biological sample to aid in the  
 CC diagnosis of HIV infection or disease status or as measure of response to  
 CC immunization.

CC Sequence 847 AA;

Query Match 88.3%; Score 2764; DB 21; Length 847;  
 Best Local Similarity 81.7%; Pred. No. 6,28-128;  
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;

```

Qy 1 VEKLVTVVYGVVPWKRAATTLTFCASDAKAYDTEVHNVAATHACVPDPDPOEVLENTV 60
Db 30 VEKLVTVVYGVVPWKRAATTLTFCASDAKAYDTEVHNVAATHACVPDPDPOEVLENTV 89
Qy 61 EHPNMKNKNWEQKQEDIIISLMDQSLKPCVLTPLCGA----- 98
Db 90 ENFNMKNKNWEQKQEDIIISLMDQSLKPCVLTPLCTYHNTNKNANTNTKSNMKEMDR 149
Qy 99 -----GDTSVITQACPKISFE 115
Db 150 GEIKNSCFKVTTSIRNKQKEYALFYKLDVVPIDNDNTSYKLINCNSTSVITQACPKISFE 209
Qy 116 PIPHYCAPAGFALCKNDKTFNGKSPCKNASTVQCTGIRRVSTQLLNGLSLABEEVY 175
Db 210 PIPHYCAPAGFALCKNDKTFNGKSPCKNASTVQCTGIRRVSTQLLNGLSLABEEVY 269
Qy 176 IRSNFTNNAKTIIVQLKESVEINCTRPNNTNRKSIHIGPGAFYTTGEIIGDIOAHCHN 235
Db 270 IRSNFTNNAKTIIVQLKESVEINCTRPNNTNRKSIHIGPGAFYTTGEIIGDIOAHCHN 329
Qy 236 ISRAKNDTLKOIVIKLEBQENKTIIVFNNSSGDPPEIVMHSFNCGEFFYCNSTOLFNS 295
Db 330 ISGKMNNTLKOIVTKLOAFQGNKTIIVFQSSGDPPEIVMHSFNCGEFFYCNSTOLFNS 389
Qy 296 TWNNNTGSSNNTGENTITLPCRKOIINMMEVQKAMTAPPIRGQIRGSSNTGLLTRD 355
Db 390 TW-NNTIGPNNTNG-TITLPCRKOIINMMEVQKAMTAPPIRGQIRGSSNTGLLTRD 447
Qy 356 GGIN-ENGTEIFRPGGDMRDNRSELYKYRVVIEPLGVAPTKRRRVQEKRAVIG 414
Db 448 GGEISNTTEIFRPGGDMRDNRSELYKYRVVIEPLGVAPTKRRRVQEKRAVIG 507
Qy 415 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQORMLQTVWGIKQL 474
Db 508 AMFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQORMLQTVWGIKQL 567
Qy 475 QARVLAVERVYLDQQLGIWGGSGKLICTTAVPNNASWSNKSLEDIRIMNMNTMWEWEREID 534
Db 568 QARVLAVERVYLDQQLGIWGGSGKLICTTAVPNNASWSNKSLEDIRIMNMNTMWEWEREID 627
Qy 535 NYTSEIYTLIESQNOQEKNEQELLELDKMASLNMWPDITMWMY 579
Db 628 NYTSEIYTLIESQNOQEKNEQELLELDKMASLNMWPDITMWMY 672

```

## RESULT 11

```

ABU66565
ID ABU66565 standard; Protein; 842 AA.
XX
AC ABU66565;
XX
DT 22-MAY-2003 (first entry)

```

```

XX DE Human immunodeficiency virus (HIV) envelope (env) protein #1.
XX XX
XX XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
XX KM vaccine; gene therapy; packaging cell line; humoral immune response;
XX KM cellular immune response; gene delivery vector; DNA immunisation;
XX XX envelope protein; env.
XX OS Human immunodeficiency virus.
XX PN WO2003004657-A1.
XX PD 16-JAN-2003.
XX PF 05-JUL-2002; 2002WO-US21421.
XX XX
XX PR 05-JUL-2001; 2001US-303192P.
XX PR 31-AUG-2001; 2001US-316860P.
XX PR 16-JAN-2002; 2002US-349728P.
XX PR 16-JAN-2002; 2002US-349793P.
XX PR 16-JAN-2002; 2002US-349871P.
XX PA (CHIR ) CHIRON CORP.
XX PI Zur Megede J, Barnett SW, Lian Y;
XX DR WPI; 2003-221602/21.
XX XX
XX PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
XX PT C polypeptides, useful as immunogenic compositions or vaccines for
XX PT generating humoral or cellular immune responses against HIV in a
XX PT subject, especially humans -
XX PS Disclosure; Fig 2A-C; 262pp; English.
XX XX
XX CC The invention describes a synthetic polynucleotide encoding 2 or more
XX CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
XX CC derived from different HIV subtypes. The polynucleotide is useful for
XX CC immunisation, generation of packaging cell lines, or production of HIV
XX CC polypeptides. The polynucleotide and its encoded proteins are useful as
XX CC immunogenic compositions or vaccines for generating humoral or cellular
XX CC immune responses against HIV in a subject, or for inducing neutralising
XX CC antibodies against HIV. The gene delivery vector comprising the
XX CC polynucleotide is also useful for DNA immunisation of, or for
XX CC generating an immune response (e.g. a humoral or cellular immune
XX CC response) in, a subject such as a mammal, particularly a human. This
XX CC is the amino acid sequence of a human immunodeficiency virus (HIV)
XX CC envelope (env) protein.
XX CC
XX CC Sequence 842 AA;
XX CC
XX CC Query Match 88.2%; Score 2759; DB 24; Length 842;
XX CC Best Local Similarity 81.6%; Pred. No. 1.1e-127;
XX CC Matches 526; Conservative 16; Mismatches 35; Indels 68; Gaps 4;
XX CC
XX QY 1 VEKLVWTVYGVPMWEKATTTTFCASDAKAYTEVNVATACVPTDNPQEVLEANT 60
XX DB 25 VEKLVWTVYGVPMWEKATTTTFCASDAKAYTEVNVATACVPTDNPQEVLEANT 84
XX QY 61 EHFNNMKNMNVQOMEDIISLWDQSLKPCVKLTPLCGA----- 98
XX DB 85 EHFNNMKNMNVQOMEDIISLWDQSLKPCVKLTPLCVLHCTNLKNAATWTKSSNMKEMDR 144
XX QY 99 -----GCDTSYITQACPKLSE 115
XX DB 145 GEIKNCSEKVTTSIRNKQKEVALFYKLDVVIDNDNTSYKLLINCNTSVITQACPKVSFE 204
XX QY 116 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTHGIRPVSTQLNLGSLAEEVY 175
XX DB 205 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTHGIRPVSTQLNLGSLAEEVY 264
XX QY 176 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGRIIDIRQAHCN 235
XX

```

```

DB 265 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYATGDIIDIRQAHCN 324
QY 236 ISRAKMNITLKOIVKLEQFENKTIIVPHSSGSDPEIVMHSFNGSGEFYCNSTQLFNS 295
DB 325 ISGERKMNITLKOIVKLEQFENKTIIVPHSSGSDPEIVMHSFNGSGEFYCNSTQLFNS 384
QY 236 TNNNTTEGSGNTEGNTITLPCRKOIIMMWOEVGKAMYAPPIRGOIRCSSNITGLLITRD 355
DB 385 TW-NTTIGPNNNG--TITLPCRKOIINRMQGVGKAMYAPPIRGOIRCSSNITGLLITRD 442
QY 356 GGIN-ENGTEIFRPGCGDMRDNRSELYKVKVKTIEPLGAVPTCKRRVQREKRAVGIG 414
DB 443 GGEKISNTTEIFRPGCGDMRDNRSELYKVKVKTIEPLGAVPTCKRRVQREKRAVGIG 502
QY 415 AVFLGPIGAAGSTMAAAMTLTVQARLLSGIVQOONNLTPRIEAOQRMLOITWAGIKQL 474
DB 503 AMFLGPIGAAGSTMAAAMTLTVQARLLSGIVQOONNLTPRIEAOQRMLOITWAGIKQL 562
QY 475 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASNSKSLDRIIMNMNTMWEWEREID 534
DB 563 QARVLAVERYLGDQQLLGWCGSGKLICTTAVPNNASNSKSLDRIIMNMNTMWEWEREID 622
QY 535 NTSRIYTLIBESQNOQEKNEQELIEDKMASLNNWFDITNNKLMY 579
DB 623 NYTNLIYTLIBESQNOQEKNEQELIEDKMASLNNWFDITNNKLMY 667
XX
XX RESULT 12
XX ID ABB06211 standard; Protein; 842 AA.
XX XX
XX AC ABB06211;
XX XX
XX DT 15-MAY-2002 (first entry)
XX XX
XX DE HIV Env isolate SF162 amino acid sequence.
XX XX
XX KM Human immunodeficiency virus type C; antigenic HIV type C protein;
XX KM immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX KM immunostimulant; gene therapy.
XX XX
XX OS Human immunodeficiency virus type C.
XX OS Synthetic.
XX PN WO200204493-A2.
XX PD 17-JAN-2002.
XX PF 05-JUL-2001; 2001WO-US21241.
XX PR 05-JUL-2000; 2000US-0610313.
XX PA (CHIR ) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX DB WPI; 2002-154920/20.
XX XX
XX PT New polynucleotides encoding antigenic HIV type C polypeptides, useful
XX PT in applications including DNA immunization or generation of packaging
XX PT cell lines, particularly in gene therapy -
XX PS Disclosure; Fig 105; 233pp; English.
XX XX
XX CC The present invention describes expression cassettes comprising a
XX CC polynucleotide sequence encoding a polypeptide comprising immunogenic
XX CC HIV type C polypeptides. The expression cassettes comprising any of the
XX CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX CC Nef (1). (1) have immunostimulant activity and can be used in gene
XX CC therapy. The HIV type C polynucleotides are useful in applications
XX CC including DNA immunisation, generation of packaging cell lines, and
XX CC production of HIV type C proteins. The polynucleotides are particularly

```

CC useful in gene therapy and DNA immunisation applications. ABL39942 to  
CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the  
CC exemplification of the present invention.

XX Sequence 842 AA;

Query Match 87.9%; Score 2749; DB 23; Length 842;  
Best Local Similarity 81.2%; Pred. No. 3.3e-127;  
Matches 524; Conservative 16; Mismatches 37; Indels 68; Gaps 4;

```

QY 1 VEKLVVTVYGVVWPKKATTTLLFCASDAKAYDEVHANWATHACVPTDPNPOEVVLENTV 60
DB 25 VEKLVVTVYGVVWPKKATTTLLFCASDAKAYDEVHANWATHACVPTDPNPOEVVLENTV 84
QY 61 EHFNMKNMVEQOQEDIIISLMDOSLKPCKVLTPLCGA----- 98
DB 85 EHFNMKNMVEQOQEDIIISLMDOSLKPCKVLTPLCGA----- 144
QY 99 -----GCDTSVITQACPKISFE 115
DB 145 GEIKNCSFKVTTISIRNKMQKEYALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKISFE 204
QY 116 PIPHYCAPAGFALIKCNDKTPNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEEV 175
DB 205 PIPHYCAPAGFALIKCNDKTPNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEEV 264
QY 176 IRSDFNTNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQACHN 235
DB 265 IRSDFNTNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQACHN 324
QY 236 ISRAKMDTLKOIYIKLREOFENKTIIVNHSFGDPEIVMHSFNGGGEFFYCNSDTQLFNS 295
DB 325 ISGEKMNNTLKOIYIKLQAGFNGKTIIVKQSSGGDEPEIVMHSFNGGGEFFYCNSDTQLFNS 384
QY 296 TWANNNTBSNNTTEGNTITLPCRIKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTRD 355
DB 385 TW-NTTIGPNNTG-TITLPCRIKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTRD 442
QY 356 GGIN-ENGTEIFRPGGDMRDMNRSELKYKVKIETPLGVAPTCKRRVVOREKRAVGIG 414
DB 443 GGEKISNTTEIFRPGGDMRDMNRSELKYKVKIETPLGVAPTCKRRVVOREKRAVGIG 502
QY 415 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQMLQTLTWGIKQL 474
DB 503 AMFLGFLGAAGSTMGASMTLTVQARQLLSGIVQOQNLLRAIEAQOQMLQTLTWGIKQL 562
QY 475 QARVLAVERYLGDQOLLGIGWCSGKLICTTAVPWNASWSNKSIDRIANNMTWMEWEREID 534
DB 563 QARVLAVERYLKDQOLLGIGWCSGKLICTTAVPWNASWSNKSIDRIANNMTWMEWEREID 622
QY 535 NYTSEIYTLIESQOQEQEKEOELELDPMASLMMWPDITNMLWY 579
DB 623 NYTNLTITLIESQOQEQEKEOELELDPMASLMMWPDITNMLWY 667

RESULT 13
AAP80967
ID AAP80967 standard; protein; 851 AA.
AC AAP80967;
XX
XX 25-MAR-2003 (updated)
XX 15-NOV-1990 (first entry)
XX
XX HIV protein HT6.
XX HIV; HT6; gp160; envelope protein; RF; AIDS.
XX OS Human immunodeficiency virus variant RF.
XX EP272858-A.
XX
XX 29-JUN-1988.

```

```

XX 14-DEC-1987; 87EP-0310967.
XX
XX 31-AUG-1987; 87US-0091481.
XX 15-DEC-1986; 86US-0941111.
XX
XX (REPK) REPLIGEN CORP.
XX
XX Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
XX WPI; 1988-176944/26.
XX DR N-PSDB; AAN80948.
XX
XX The sequence is the result of cloning a hybrid envelope gene from
XX HIV variants BH10 and RF. A central portion of the RF gene was
XX used, the rest being from the distantly related variant BH10. The
XX resulting clone, pACH6 produces a hybrid gp 160 envelope protein
XX with novel immunological and antigenic characteristics. It may
XX be used to as a vaccine and for diagnosis and therapy of AIDS.
XX See also AAP80966.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 851 AA;
XX
XX Query Match 87.4%; Score 2735.5; DB 9; Length 851;
XX Best Local Similarity 79.9%; Pred. No. 1.5e-126;
XX Matches 516; Conservative 33; Mismatches 29; Indels 69; Gaps 4;

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHANWATHACVPTDPNPOEVVLENTV 61
DB 32 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHANWATHACVPTDPNPOEVVLENTV 91
QY 62 HFNMMKNMVEQOQEDIIISLMDOSLKPCKVLTPLCGA----- 98
DB 92 HFNMMKNMVEQOQEDIIISLMDOSLKPCKVLTPLCGA----- 151
QY 99 -----GCDTSVITQACPKISFE 115
DB 152 GEIKNCSFNNSTSIKQKEYAFYKLDIIPIDNDTSTYTLNCSNTSVITQACPKISFE 211
QY 116 PIPHYCAPAGFALIKCNDKTPNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEEV 175
DB 212 PIPHYCAPAGFALIKCNDKTPNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEEV 271
QY 176 IRSDFNTNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQACHN 235
DB 272 IRSDFNTNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQACHN 331
QY 236 ISRAKMDTLKOIYIKLREOFENKTIIVNHSFGDPEIVMHSFNGGGEFFYCNSDTQLFNS 295
DB 332 ISRAKMDTLKOIYIKLREOFENKTIIVNHSFGDPEIVMHSFNGGGEFFYCNSDTQLFNS 391
QY 296 TWANNNTBSNNTTEGNTITLPCRIKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTR 354
DB 392 TW-NTTIGPNNTG-TITLPCRIKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTR 450
QY 355 DGIN-ENGTEIFRPGGDMRDMNRSELKYKVKIETPLGVAPTCKRRVVOREKRAVGIG 413
DB 451 DGEDTNTTEIFRPGGDMRDMNRSELKYKVKIETPLGVAPTCKRRVVOREKRAVGIG 510
QY 414 GAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQMLQTLTWGIKQL 473
DB 511 GALFLGFLGAAGSTMGASMTLTVQARQLLSGIVQOQNLLRAIEAQOQMLQTLTWGIKQL 570
QY 474 LQARVLAVERYLGDQOLLGIGWCSGKLICTTAVPWNASWSNKSIDRIANNMTWMEWERE 533
DB 571 LQARVLAVERYLKDQOLLGIGWCSGKLICTTAVPWNASWSNKSIDRIANNMTWMEWERE 630

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:28:04 ; Search time 13.1662 Seconds  
(without alignments)  
1860.670 Million cell updates/sec

Title: US-10-032-162-15

Sequence: 1 VEKLVTVYGVGVKWEAT.....ELDKWASLWNPFDITWLMY 579

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/6C\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/6D\_COMB.pep:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2800.5	89.5	855	1 US-08-022-835-6	Sequence 6, Appl
2	2800.5	89.5	855	1 US-08-388-809-6	Sequence 6, Appl
3	2800.5	89.5	855	1 US-08-647-714-6	Sequence 6, Appl
4	2776.5	88.7	850	2 US-08-448-603A-28	Sequence 28, Appl
5	2776.5	88.7	850	2 US-09-134-075-28	Sequence 28, Appl
6	2776.5	88.7	850	4 US-09-492-739-28	Sequence 28, Appl
7	2754.5	88.0	855	3 US-07-956-483-14	Sequence 14, Appl
8	2725	87.1	857	2 US-08-448-603A-30	Sequence 30, Appl
9	2725	87.1	857	2 US-09-134-075-30	Sequence 30, Appl
10	2725	87.1	857	4 US-09-492-739-30	Sequence 30, Appl
11	2724.5	87.1	861	1 US-08-127-499A-14	Sequence 14, Appl
12	2724.5	87.1	861	1 US-08-482-867-14	Sequence 14, Appl
13	2724.5	87.1	861	1 US-07-956-483-10	Sequence 10, Appl
14	2724.5	87.1	861	3 US-08-472-240A-7	Sequence 7, Appl
15	2724.5	87.1	861	3 US-08-472-240A-7	Sequence 7, Appl
16	2724.5	87.1	861	4 US-08-817-441-103	Sequence 103, Appl
17	2719	86.9	856	3 US-09-124-900-9	Sequence 9, Appl
18	2719	86.9	856	3 US-08-463-210-11	Sequence 11, Appl
19	2719	86.9	856	3 US-08-472-240A-5	Sequence 5, Appl
20	2716.5	86.8	861	3 US-07-956-483-16	Sequence 16, Appl
21	2712	86.7	826	1 US-08-375-510-2	Sequence 2, Appl
22	2712	86.7	826	2 US-08-487-657-2	Sequence 2, Appl
23	2707.5	86.5	854	4 US-09-309-572-23	Sequence 23, Appl
24	2707.5	86.5	854	4 US-08-788-815-7	Sequence 7, Appl
25	2706	86.5	855	3 US-09-157-963-7	Sequence 7, Appl
26	2706	86.5	855	3 US-07-956-483-15	Sequence 15, Appl
27	2702	86.4	856	2 US-07-916-098A-2	Sequence 2, Appl

28	2695.5	86.1	839	3 US-08-472-240A-10	Sequence 10, Appl
29	2694	86.1	856	4 US-09-337-387-11	Sequence 11, Appl
30	2690	86.0	856	1 US-08-022-835-2	Sequence 2, Appl
31	2690	86.0	856	1 US-08-388-809-2	Sequence 2, Appl
32	2690	86.0	856	2 US-08-647-714-2	Sequence 2, Appl
33	2690	86.0	856	3 US-07-956-483-11	Sequence 11, Appl
34	2686	85.8	857	1 US-08-022-835-4	Sequence 4, Appl
35	2686	85.8	857	1 US-08-388-809-4	Sequence 4, Appl
36	2686	85.8	857	2 US-08-647-714-4	Sequence 4, Appl
37	2672.5	85.4	726	4 US-09-337-387-3	Sequence 3, Appl
38	2670	85.3	887	3 US-08-472-240A-6	Sequence 6, Appl
39	2663	85.1	865	3 US-07-956-483-13	Sequence 13, Appl
40	2654	84.8	856	1 US-08-375-510-1	Sequence 1, Appl
41	2651.5	84.7	759	4 US-09-337-387-12	Sequence 12, Appl
42	2648.5	84.6	657	4 US-09-256-194-2	Sequence 2, Appl
43	2639	84.3	887	3 US-08-472-240A-2	Sequence 2, Appl
44	2626	83.9	887	3 US-08-472-240A-4	Sequence 4, Appl
45	2620.5	83.7	665	2 US-08-448-603A-32	Sequence 32, Appl

# ALIGNMENTS

RESULT 1  
US-08-022-835-6  
; Sequence 6, Application US/08022835  
; Patent No. 5420030  
; GENERAL INFORMATION:  
; APPLICANT: Reitz Jr., Marvin S.  
; APPLICANT: Franchini, Genoveffa  
; APPLICANT: Markham, Phillip D.  
; APPLICANT: Gallo, Robert C.  
; APPLICANT: Lori, Franco C.  
; APPLICANT: Popovics, Mikulas  
; APPLICANT: Garner, Suzanne  
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: Eleventh Floor, 1615 L. Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/022,835  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/599,491  
FILING DATE: 17-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-022-835-6

Query Match 89.5%; Score 2800.5; DB 1; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 1.5e-233;  
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;

QY 2 EKLWVTVYGVVWKEATTTTLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNTE 61  
 DB 31 EKLWVTVYGVVWKEATTTTLFCASDRKAYDEVHANNWATHACVPTDPNPOEVLNTE 90  
 QY 62 HNNMKNMVEQWOMEDIISLMDOSLKPCVKLTPLC----- 96  
 DB 91 HNNMKNMVEQWOMEDIISLMDOSLKPCVKLTPLC----- 150  
 QY 97 --GAG-----CDSVTITQACP 110  
 DB 151 MVGGEMKNCSFNITTNIRGVQKEVALFYKLDIAPIDNNSNNRYLISCTSVITQACP 210  
 QY 111 KISFEPIPIHCAAGPAILKCNKTPNGKPCGVSTVQCTHGRPVSTQOLLNGLA 170  
 DB 211 KVSFEPIPIHCAAGPAILKCKDKKFKNGKPCGVSTVQCTHGRPVSTQOLLNGLA 270  
 171 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 230  
 DB 271 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 330  
 QY 231 QAHCNISRAKNDTLKOIVIKLREQFNKTIIVFNHSSGDEIIVHSPFCGGEFFYCNS 290  
 DB 331 QAHCNLSRAKNDTLKIVIKLREQFNKTIIVFKHSSGDEIIVHSPFCGGEFFYCNS 390  
 QY 291 QLFNSTWNNTEGSSNNTGNTTLPCKRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 350  
 DB 391 QLFNSTWNNTEGSSNNTGNTTLPCKRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 450  
 QY 351 LLTRDGGINGETIEFRPGGDMRDNRSELYKKVYKIEPLGVAFTPKRRVQREKRA 410  
 DB 451 LLTRDGGEDKTEIEFRPGGDMRDNRSELYKKVYKIEPLGVAFTPKRRVQREKRA 510  
 QY 411 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQLTLTWG 470  
 DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQLTLTWG 570  
 QY 471 IKOLQARVLAVERLYGDOOLIGIGCSGKLTCTTAVPNNASWSNKSIDRIINNNTWMEWE 530  
 DB 571 IKOLQARVLAVERLYGDOOLIGIGCSGKLTCTTAVPNNASWSNKSIDRIINNNTWMEWE 630  
 QY 531 REIDNYSIETLIEESONOEKNEOELELDKASLNMWPDITWMLWY 579  
 DB 631 REIDNYSIETLIEESONOEKNEOELELDKASLNMWPDITWMLWY 679

US-08-388-809-6  
 Sequence 6, Application US/08388809  
 Patent No. 5576000  
 GENERAL INFORMATION:  
 APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
 APPLICANT: GENOVEFA; MARKHAM, PHILIP D.; GALLO, ROBERT  
 APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND  
 APPLICANT: CARTNER, SUZANNE  
 TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
 TITLE OF INVENTION: AND USES THEREOF  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORGAN & PINNEGAN  
 STREET: 345 PARK AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10154  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY DISK, 3.5"  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/388, 809  
 FILING DATE: 15-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/022, 835  
 FILING DATE: 25-FEB-1993  
 APPLICATION NUMBER: US 07/599, 491  
 FILING DATE: 17-OCT-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LESLIE A. SERUNIAN  
 REGISTRATION NUMBER: 35,353  
 REFERENCE/DOCKET NUMBER: 2026-4092052  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 758-4800  
 TELEFAX: (212) 751-6849  
 TELEX: 421792  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 855 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-388-809-6

Query Match 89.5%; Score 2800.5; DB 1; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 1.5e-233;  
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;

QY 2 EKLWVTVYGVVWKEATTTTLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNTE 61  
 DB 31 EKLWVTVYGVVWKEATTTTLFCASDRKAYDEVHANNWATHACVPTDPNPOEVLNTE 90  
 QY 62 HNNMKNMVEQWOMEDIISLMDOSLKPCVKLTPLC----- 96  
 DB 91 HNNMKNMVEQWOMEDIISLMDOSLKPCVKLTPLC----- 150  
 QY 97 --GAG-----CDSVTITQACP 110  
 DB 151 MVGGEMKNCSFNITTNIRGVQKEVALFYKLDIAPIDNNSNNRYLISCTSVITQACP 210  
 QY 111 KISFEPIPIHCAAGPAILKCNKTPNGKPCGVSTVQCTHGRPVSTQOLLNGLA 170  
 DB 211 KVSFEPIPIHCAAGPAILKCKDKKFKNGKPCGVSTVQCTHGRPVSTQOLLNGLA 270  
 171 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 230  
 DB 271 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 330  
 QY 231 QAHCNISRAKNDTLKOIVIKLREQFNKTIIVFNHSSGDEIIVHSPFCGGEFFYCNS 290  
 DB 331 QAHCNLSRAKNDTLKIVIKLREQFNKTIIVFKHSSGDEIIVHSPFCGGEFFYCNS 390  
 QY 291 QLFNSTWNNTEGSSNNTGNTTLPCKRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 350  
 DB 391 QLFNSTWNNTEGSSNNTGNTTLPCKRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 450  
 QY 351 LLTRDGGINGETIEFRPGGDMRDNRSELYKKVYKIEPLGVAFTPKRRVQREKRA 410  
 DB 451 LLTRDGGEDKTEIEFRPGGDMRDNRSELYKKVYKIEPLGVAFTPKRRVQREKRA 510  
 QY 411 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQLTLTWG 470  
 DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQLTLTWG 570  
 QY 471 IKOLQARVLAVERLYGDOOLIGIGCSGKLTCTTAVPNNASWSNKSIDRIINNNTWMEWE 530  
 DB 571 IKOLQARVLAVERLYGDOOLIGIGCSGKLTCTTAVPNNASWSNKSIDRIINNNTWMEWE 630  
 QY 531 REIDNYSIETLIEESONOEKNEOELELDKASLNMWPDITWMLWY 579  
 DB 631 REIDNYSIETLIEESONOEKNEOELELDKASLNMWPDITWMLWY 679

RESULT 3  
US-08-647-714-6  
Sequence 6, Application US/08647714  
Patent No. 5869313  
GENERAL INFORMATION:  
APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,  
APPLICANT: GENOVEFA; MARKAM, PHILIP D. GALLO, ROBERT  
APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND  
APPLICANT: GARTNER, SUZANNE  
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: MORGAN & PINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK, 3.5"  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/647,714  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,809  
FILING DATE: 15-FEB-1995  
APPLICATION NUMBER: US 08/022,835  
FILING DATE: 25-FEB-1993  
APPLICATION NUMBER: US 07/599,491  
FILING DATE: 17-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: LESLIE A. SERUNIAN  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4092US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DB-647-714-6

Query Match 89.5%; Score 2800.5; DB 2; Length 855;  
Best Local Similarity 81.2%; Pred. No. 1.5e-233;  
Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;

QY 2 EKLWTVVYGVVWVKEATTTTLCASDAKAYDEVANWVATHACVPTDPNPOEVLNTE 61  
DB 31 EKLWTVVYGVVWVKEATTTTLCASDAKAYDEVANWVATHACVPTDPNPOEVLNTE 90  
QY 62 HFNWAKNNWVWQMODIISLWQSLKPCYKLTPLC----- 96  
DB 91 NFNWAKNNWVWQMODIISLWQSLKPCYKLTPLC----- 150  
QY 97 --GAG-----CDSVITQCP 110  
DB 151 WVGGEKMKVCSNITNTIRGKQKAYALFYKLDIAPIDNNSNNRYRLISCNTSVITQCP 210  
QY 111 KISEPPIPIHYCAPAGFALILKCDNKTENGKGPCKNVSTVOCTGIRPVVSTOLLNGSLA 170  
DB 211 KVSFPIPIHYCAPAGFALILKCDKFKNGKPCNVSTVOCTGIRPVVSTOLLNGSLA 270  
QY 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGRAPFTTGEIIGDIR 230  
DB 271 EEEVIRSANPADNAKVIIVOLNSESVEINCTRPNNNTRKSHIGGRAPFTTGEIIGDIR 330

QY 231 QAHCNISRAKNWDLTKQIVIKLREQENKTIIVFNHSGGDEIVNHSFNCGEFFYCNS 290  
DB 331 QAHCNISRAKNWDLTKQIVIKLREQENKTIIVFNHSGGDEIVNHSFNCGEFFYCNS 390  
QY 291 QLFNSTNNNTEGNSNTEGNTITTLPCRIKQIIMWQEVKAMYPPIRGQIRCSNITGL 350  
DB 391 QLFNSTNNNTEGNSNTEGNTITTLPCRIKQIIMWQEVKAMYPPIRGQIRCSNITGL 450  
QY 351 LITRGGINGNGETIFRGGGDMRNMRSELYKVVVIEPLGVAPTCKRVRVQREKA 410  
DB 451 LITRGGINGNGETIFRGGGDMRNMRSELYKVVVIEPLGVAPTCKRVRVQREKA 510  
QY 411 VGIQAVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIBAOGRMLQTYWG 470  
DB 511 VGIQAVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIBAOGRMLQTYWG 570  
QY 471 IKQLQARVLAVERVYIGDOOLLGIGWCSGKLICTTAVPNNASWSNKSILDRINNMWME 530  
DB 571 IKQLQARVLAVERVYIGDOOLLGIGWCSGKLICTTAVPNNASWSNKSILDRINNMWME 630  
QY 531 REINDYTSIYTLIESONQOEKNEQLLELDKASLNNWPDITWMLY 579  
DB 631 REINDYTSIYTLIESONQOEKNEQLLELDKASLNNWPDITWMLY 679

RESULT 4  
US-08-448-603A-28  
Sequence 28, Application US/08448603A  
Patent No. 5864027  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
APPLICANT: Nakamura, Gerald R.  
TITLE OF INVENTION: HIV Envelope Polypeptides  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESS: McCutchen, Doyle, Brown & Enersen, LLP  
STREET: 3 Embarcadero Center  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,603A  
FILING DATE: 07-JUN-1994  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/072,833  
FILING DATE: 07-JUN-93  
ATTORNEY/AGENT INFORMATION:  
NAME: Haliday, Emily  
REGISTRATION NUMBER: 38903  
REFERENCE/DOCKET NUMBER: 14918-704  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-393-2000  
TELEFAX: 415-393-2286  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 850 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-448-603A-28

Query Match 88.7%; Score 2776.5; DB 2; Length 850;  
Best Local Similarity 81.5%; Pred. No. 1.7e-231;  
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

```

QY 2 EKLWVTVYGVVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWVTVYGVVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNMVMQOMEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NFNWKNMVMQOMEDIIISLMDQSLKPCVKLTPLCVLTNCTDLKNAATNTSSWGKMERG 150
QY 99 -----GDDTSVITQACPKISFEP 116
DB 151 EIKNCSFNVTTSIRDKMKNKNEVALFYKLDVVDIDNDNTSYRLISCTSVITQACPKISFEP 210
QY 117 IPIHYCAPAGPAILKCNKDKTEFNKGPKCNKAVSTVOCTHGIRPVVSTOLLNGSLAESEVVI 176
DB 211 IPIHYCAPAGPAILKCNKDKTEFNKGPKCNKAVSTVOCTHGIRPVVSTOLLNGSLAESEVVI 270
QY 177 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHONI 236
DB 271 RSNFSDNAKTIIVQLNBSVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
QY 237 SRAMNDTLKQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNGGFEFFCNSTOLFNST 296
DB 331 SSTRMNTLQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNGGFEFFCNSTOLFNST 389
QY 297 WN-----NTEGSNNTGNTITLPCRKOIINMNOEKGKAMVAPPIRGQIRCSNITGL 351
DB 390 WNYTYTNNTGSDNT-GRNITLQCRKOIINMNOEKGKAMVAPPIRGQIRCSNITGL 448
QY 352 LTRDGINENGTETFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRKRVQREKRAV 411
DB 449 LTRDGG-NNSETEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRKRVQREKRAV 507
QY 412 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEBOHLLQLTWGI 471
DB 508 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEBOHLLQLTWGI 567
QY 472 KQLOARVLAVERYLKQDQDLGIMGSGKLTCTTAVPNAWSNKSILDRINNMTMMEWER 531
DB 568 KQLOARVLAVERYLKQDQDLGIMGSGKLTCTTAVPNAWSNKSILDRINNMTMMEWER 627
QY 532 EIDNVTSEIYTLIEESQNOEKNEQELLELDKMASLNNWFDITWMLWY 579
DB 628 EIDNVTSEIYTLIEESQNOEKNEQELLELDKMASLNNWFDITWMLWY 675

```

RESULT 5  
 US-09-134-075-28  
 Sequence 28, Application US/09134075  
 Patent No. 6042836  
 GENERAL INFORMATION:  
 APPLICANT: Berman, Phillip W.  
 APPLICANT: Nakamura, Gerald R.  
 TITLE OF INVENTION: HIV Envelope Polypeptides  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McCutchen, Doyle, Brown & Eversen, LLP  
 STREET: 3 Embarcadero Center  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/134,075  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/448,603

```

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-134-075-28
Query Match 88.7%; Score 2776.5; DB 3; Length 850;
Best Local Similarity 81.5%; Pred. No. 1,7e-231;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

```

```

QY 2 EKLWVTVYGVVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWVTVYGVVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNMVMQOMEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NFNWKNMVMQOMEDIIISLMDQSLKPCVKLTPLCVLTNCTDLKNAATNTSSWGKMERG 150
QY 99 -----GDDTSVITQACPKISFEP 116
DB 151 EIKNCSFNVTTSIRDKMKNKNEVALFYKLDVVDIDNDNTSYRLISCTSVITQACPKISFEP 210
QY 117 IPIHYCAPAGPAILKCNKDKTEFNKGPKCNKAVSTVOCTHGIRPVVSTOLLNGSLAESEVVI 176
DB 211 IPIHYCAPAGPAILKCNKDKTEFNKGPKCNKAVSTVOCTHGIRPVVSTOLLNGSLAESEVVI 270
QY 177 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHONI 236
DB 271 RSNFSDNAKTIIVQLNBSVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
QY 237 SRAMNDTLKQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNGGFEFFCNSTOLFNST 296
DB 331 SSTRMNTLQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNGGFEFFCNSTOLFNST 389
QY 297 WN-----NTEGSNNTGNTITLPCRKOIINMNOEKGKAMVAPPIRGQIRCSNITGL 351
DB 390 WNYTYTNNTGSDNT-GRNITLQCRKOIINMNOEKGKAMVAPPIRGQIRCSNITGL 448
QY 449 LTRDGG-NNSETEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRKRVQREKRAV 507
QY 412 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEBOHLLQLTWGI 471
DB 508 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEBOHLLQLTWGI 567
QY 472 KQLOARVLAVERYLKQDQDLGIMGSGKLTCTTAVPNAWSNKSILDRINNMTMMEWER 531
DB 568 KQLOARVLAVERYLKQDQDLGIMGSGKLTCTTAVPNAWSNKSILDRINNMTMMEWER 627
QY 532 EIDNVTSEIYTLIEESQNOEKNEQELLELDKMASLNNWFDITWMLWY 579
DB 628 EIDNVTSEIYTLIEESQNOEKNEQELLELDKMASLNNWFDITWMLWY 675

```

RESULT 6  
 US-09-492-739-28  
 Sequence 28, Application US/09492739  
 Patent No. 6331404  
 GENERAL INFORMATION:  
 APPLICANT: Berman, Phillip W.  
 APPLICANT: Nakamura, Gerald R.

```

/ TITLE OF INVENTION: HIV Envelope Polypeptides
/ NUMBER OF SEQUENCES: 33
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: McCutchen, Doyle, Brown & Eneersen, LLP
/ STREET: 3 Embarcadero Center
/ CITY: San Francisco
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94111
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/492,739
/ FILING DATE: 27-Jan-2000
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/134,075
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haliday, Emily
/ REGISTRATION NUMBER: 38903
/ REFERENCE/DOCKET NUMBER: 14918-704
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-393-2000
/ TELEFAX: 415-393-2286
/ TELETYPE: <Unknown>
/ INFORMATION FOR SEQ ID NO: 28:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 850 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-492-739-28

Query Match      88.7% Score 2776.5; DB 4; Length 850;
Best Local Similarity 81.5%; Pred. No. 1.7e-231;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

QY 2 EKLWTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNWATTHACVPTDPNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNWATTHACVPTDPNPQEVLENTYE 90
QY 62 HNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 98
QY 91 NNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 150
DB 99 -----GCDTSVTTQCPKISFEP 116
DB 151 EIKNCSFNTTISIRDMKNEVALFYKLDVPIIDNDNTSYRLISCTSVITQACPKVSFEP 210
QY 117 IPIHYCAPAGFAILKCNDRKTPNGKPCKNVSTVOCTHGIRPVYSTQLLNGSLAEBEVYI 176
DB 211 IPIHYCAPAGFAILKCRDKKFGTGTCTVSTVOCTHGIRPVYSTQLLNGSLAEBEVYI 270
QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGIRQAHNI 236
DB 271 RSNANFSDNAKTIIVOLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGIRQAHNI 330
QY 237 SAKKXNDTLKOIVIKLRFENKTIIVNHSNGGDEPIVNHSGGSEFFYCNSTQLFNT 296
DB 331 SSKNNNTLKOIVIKLRFENKTIIVNHSNGGDEPIVNHSGGSEFFYCNSTQLFNT 389
QY 297 WN-----NNTGSSNNTGNTITLPCRKIOIIMMOEGVGRAMYAPIRGOIRGSSNITGL 351
DB 399 WNYTWTNNTEGSDNI-GRNITLQCRKIOIIMMOEGVGRAMYAPIRGOIRGSSNITGL 448
QY 352 LTRDGINENGTEIFRPGGDMRDNRSELVYKVKVIEPLGVAFTKCRKRRVVRQEKRAV 411
DB 449 LTRDGG-NNSETEIFRPGGDMRDNRSELVYKVKVIEPLGVAFTKCRKRRVVRQEKRAV 507

```

```

QY 412 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQLTWGI 471
DB 508 GIGAVFLGFLGAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQLTWGI 567
QY 472 KQLOARVLAVERYLADQDLGIGWCSGKLICTAVPNNASNSKSLDIANNMTMEWER 531
DB 568 KQLOARVLAVERYLADQDLGIGWCSGKLICTAVPNNASNSKSLDIANNMTMEWER 627
QY 532 BIDNTSILYSLIESQNOQEKNEBELLEKMASLWMPDITWLMY 579
DB 628 BIDNTSILYSLIESQNOQEKNEBELLEKMASLWMPDITWLMY 675

RESULT 7
US-07-956-483-14
/ Sequence 14, Application US/07956483
/ Patent No. 6261799
/ GENERAL INFORMATION:
/ APPLICANT: KIENY, Marie-Paule
/ TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
/ TITLE OF INVENTION: gp160 VARIANT
/ NUMBER OF SEQUENCES: 26
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22133-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/956,483
/ FILING DATE: 31-DEC-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO 92/19742
/ FILING DATE: 12-NOV-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 91 05392
/ FILING DATE: 02-MAY-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crane-Feury, Sharon E
/ REGISTRATION NUMBER: 36,113
/ REFERENCE/DOCKET NUMBER: 017753-005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 855 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-07-956-483-14

Query Match      88.0% Score 2754.5; DB 3; Length 855;
Best Local Similarity 80.1%; Pred. No. 1.4e-229;
Matches 523; Conservative 26; Mismatches 27; Indels 77; Gaps 6;

QY 2 EKLWTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNWATTHACVPTDPNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNWATTHACVPTDPNPQEVLENTYE 90
QY 62 HNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 150

```

```

QY 99 -----GCDTSVITQACP 113
DB 151 KMEGEMTNCNSFNITTSIRSKVQKEVALFYKLDVVPIDNITSYTLINCNSTVITQACPKVS 210
QY 114 FEPPIPIHYCAPAGAILKNDKXTNGKGPCKNVSTVCTHGIRRVVSTOLLNLSLAE 173
DB 211 FEPPIPIHYCA-RWFAIILCNKKKFNCTGTCNVSTVCTHGIRRVVSTOLLNLSLAE 269
QY 174 VVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIOAH 233
DB 270 VVLSSENFIDAKTIIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIOAH 329
QY 234 CNISRAKMDTLKOIIVIKLREOFENKTIIVFNHSSGDPPEIIVMHSFNGGFEFFCNSTOLF 293
DB 330 CNISRAKMDNTLKOIIVIKLREOFENKTIIVFNHSSGDPPEIIVMHSFNGGFEFFCNSTOLF 389
QY 294 NSTNNNTGSSNTEGN-TTILPCRIKOIIMMOEYKAMAPPIRGQIRCSSNITGILL 352
DB 390 SSTW-NGTEGSSNNTGNDITILPCRIKEIIMMOEYKAMAPPIRGQIRCSSNITGILL 448
QY 353 TROGGINENG-----TEIFRPGGDMRDMNRSELYKYVVKIEPLGVAPTCKRKYVORE 407
DB 449 TRDGSNSNGSKNENTEIFRPGGDMRDMNRSELYKYVVKIEPLGVAPTAKRRVYORE 508
QY 408 KRAVG-IGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNLLRAIEAOQRLQL 466
DB 509 KRAVGTIGAMFLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNLLRAIEAOQRLQL 568
QY 467 TWGKIQOLARVLAVERYLDQOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTW 526
DB 569 TWGKIQOLARVLAVERYLDQOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTW 628
QY 527 MEMREIDNYSSEIYTLIESONOEKNEOELELDKASILMNPEDITWMLMY 579
DB 629 MEMREIDNYSIYTLIESONOEKNEOELELDKASILMNPEDITWMLMY 681

```

## RESULT 8

```

US-08-448-603A-30
; Sequence 30, Application US/08448603A
; Patent No. 3864027

```

```

; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASCAL for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448.603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/072.833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:

```

```

; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-448-603A-30

```

```

Query Match 87.1%; Score 2725; DB 2; Length 857;
Best Local Similarity 78.6%; Pred. No. 5e-227;
Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

```

QY 3 KLVWTVYGVVPMKEATTLTFCASDAKAYDEVANWATHACVPTDNPQEVLENVTEH 62
DB 32 KLVWTVYGVVPMKEATTLTFCASDAKAYDEVANWATHACVPTDNPQEVLENVTEH 91
QY 63 FNNKKNMNVQOMEDIISLMDQSLKPCVKLTPLC-----GAG----- 99
DB 92 FNNKKNMNVQOMEDIISLMDQSLKPCVKLTPLCVTLNCTDAQNTTNTSSREKLEKE 151
QY 100 -----GDSVITQACP 110
DB 152 IKNCSEFNTTSVADKXOKETALFNKLDIVPIDDDDRSTNSTNYRLISCNSTVITQACP 211
QY 111 KISFEPPIHYCAPAGAILKNDKXTNGKGPCKNVSTVCTHGIRRVVSTOLLNLSLA 170
DB 212 KISFEPPIHYCAPAGAILKNDKXTNGKGPCKNVSTVCTHGIRRVVSTOLLNLSLA 271
QY 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIR 230
DB 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIR 331
QY 231 QAHNISRANKMDTLKOIIVIKLREOFENKTIIVFNHSSGDPPEIIVMHSFNGGFEFFCNST 290
DB 332 QAHNISRANKMDTLKOIIVIKLREOFENKTIIVFNHSSGDPPEIIVMHSFNGGFEFFCNST 391
QY 392 QLFSTWNTDKVNSGTSTENSTITLPCRIKOIIMMOEYKAMAPPIRGQIRCSSNIT 451
DB 451 QLFSTWNTDKVNSGTSTENSTITLPCRIKOIIMMOEYKAMAPPIRGQIRCSSNIT 511
QY 409 RAVGIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNLLRAIEAOQRLQL 466
DB 512 RAVGIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNLLRAIEAOQRLQL 571
QY 469 WGIKIQOLARVLAVERYLDQOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTW 528
DB 572 WGIKIQOLARVLAVERYLDQOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTW 631
QY 529 MEMREIDNYSSEIYTLIESONOEKNEOELELDKASILMNPEDITWMLMY 579
DB 632 MEMREIDNYSIYTLIESONOEKNEOELELDKASILMNPEDITWMLMY 682

```

## RESULT 9

```

US-09-134-075-30
; Sequence 30, Application US/09134075
; Patent No. 6042836

```

```

; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111

```



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,075
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/448,603
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-134-075-30

```

```

Query Match      87.1%; Score 2725; DB 3; Length 857;
Best Local Similarity 78.6%; Pred. No. 5e-227;
Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

```

QY 3 KLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTENH 62
DB 32 KLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTENH 91
QY 63 FNMKNKNVQEQWEDIIISLDQSLKPCVLTPLC-----GAG----- 99
DB 92 FNMKNKNVQEQWEDIIISLDQSLKPCVLTPLC-----GAG----- 151
QY 100 ----- 110
DB 152 IKNSCFNITTSVRDKQKETALENKLDIVPIDDDRNSTRNSTNYRLISCNTSVITQACP 211
QY 111 KISFEPIPIHYCAPAGFALIKCNDKTFNGSGPCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
DB 212 KISFEPIPIHYCAPAGFALIKCNDKTFNGSGPCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 171 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 230
DB 272 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331
QY 231 QAHNCISRAKANDTLKQIYIKLREOFENKTIIVFNHSSGGDEPIVMSFNGSGEFPYCNST 290
DB 332 QAHNCISRAKANDTLKQIYIKLREOFENKTIIVFNHSSGGDEPIVMSFNGSGEFPYCNST 391
QY 291 QLFNSTNNNN--TEGSSNTEGNTITLPCIKQIIMMOEVGKAMAPPRGQIRCSSNT 348
DB 392 QLFNSTNNNN--TEGSSNTEGNTITLPCIKQIIMMOEVGKAMAPPRGQIRCSSNT 451
QY 349 GILLTRDGGINENGTEIFRPGGDMRDMNRSELKYKVKIPLGVAAPKCRARVQREK 408
DB 452 GILLTRDGGINENGTEIFRPGGDMRDMNRSELKYKVKIPLGVAAPKCRARVQREK 511
QY 409 RAVGIGAVFLGFLGAAGSTMGAASTLTVQARLLSGIVQOQNNTLRATEAQOQMLQTLV 468
DB 512 RAVGIGAVFLGFLGAAGSTMGAASTLTVQARLLSGIVQOQNNTLRATEAQOQMLQTLV 571
QY 469 WGIKLOLQVLAVERYILGDOOLIGIWCSSGKLICTAAPPNMASSNKSIDRIINNNWME 528
DB 572 WGIKLOLQVLAVERYILGDOOLIGIWCSSGKLICTAAPPNMASSNKSIDRIINNNWME 631
QY 529 WEREIDNTSEIYTLIESSQOQEKNEOBLELDKWAISLNNPDIITNMLWY 579

```

```

DB 632 WEREIDNTSEIYTLIESSQOQEKNEOBLELDKWAISLNNPDIITNMLWY 682

```

```

RESULT 10
US-09-492-739-30
Sequence 30, Application US/09492739
Patent No. 631404
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
Nakamura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: McCutchen, Doyle, Brown & Eversen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-492-739-30

```

```

Query Match      87.1%; Score 2725; DB 4; Length 857;
Best Local Similarity 78.6%; Pred. No. 5e-227;
Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

```

QY 3 KLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTENH 62
DB 32 KLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTENH 91
QY 63 FNMKNKNVQEQWEDIIISLDQSLKPCVLTPLC-----GAG----- 99
DB 92 FNMKNKNVQEQWEDIIISLDQSLKPCVLTPLC-----GAG----- 151
QY 100 ----- 110
DB 152 IKNSCFNITTSVRDKQKETALENKLDIVPIDDDRNSTRNSTNYRLISCNTSVITQACP 211
QY 111 KISFEPIPIHYCAPAGFALIKCNDKTFNGSGPCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
DB 212 KISFEPIPIHYCAPAGFALIKCNDKTFNGSGPCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 171 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 230
DB 272 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331

```

QY 231 QAHNCNISRANKNDTLKQIVIKLRQEFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYCNS 290  
 DB 332 QAHNCNISRANKNDTLKQIVIKLRQEFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYCNS 391  
 QY 291 QFNSTNNNN--TEGSNNTEGNTITLPCRIOIIMMOEVEKAMVAPPIRQIRCSNIT 348  
 DB 392 QFNSTNNNN--TEGSNNTEGNTITLPCRIOIIMMOEVEKAMVAPPIRQIRCSNIT 451  
 QY 349 GLLLTROGIIENGTEIFRPGGDMRDWRSELYKVKVTEPIGVAPTKRRVOREK 408  
 DB 452 GLLLTROGIIENGTEIFRPGGDMRDWRSELYKVKVTEPIGVAPTKRRVOREK 511  
 QY 409 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEMLOLV 468  
 DB 512 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEMLOLV 571  
 QY 469 WGIKOLQARVLAVERYLDQOLLGIGCSGKLICTTAVPMAWSNKSLEDRIMNMTWME 528  
 DB 572 WGIKOLQARVLAVERYLDQOLLGIGCSGKLICTTAVPMAWSNKSLEDRIMNMTWME 631  
 QY 529 WEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASIMNFDITNMLWY 579  
 DB 632 WEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASIMNFDITNMLWY 682

## RESULT 11

US-08-127-499A-14  
 ; Sequence 14, Application US/08127499A  
 ; Patent No. 5510264  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VAN ALSTYNE, Diane  
 ; APPLICANT: SHARMA, Lawrence Rajendra  
 ; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
 ; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
 ; NUMBER OF SEQUENCES: 40  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/127,499A  
 ; FILING DATE: 28-SEP-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 51916/102/INBI  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 672-5300  
 ; TELEFAX: (202) 672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 861 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; US-08-127-499A-14

Query Match 87.1%; Score 2724.5; DB 1; Length 861;  
 Best Local Similarity 79.6%; Pred. No. 5-5e-227;  
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;  
 QY 2 EKLWTVYVGVVWKEATTTTCASDARAYDTEVHNWATACVPTDPNPOEVLNENY 61  
 DB 32 EKLWTVYVGVVWKEATTTTCASDARAYDTEVHNWATACVPTDPNPOEVLNENY 91

QY 62 HFNWKNMTEOMQOEDIIISLMDQSLKPCVLTPLCGA----- 98  
 DB 92 HFNWKNMTEOMQOEDIIISLMDQSLKPCVLTPLCGA----- 151  
 QY 99 -----GDDTVITQACP 110  
 DB 152 MMEKEGEIKNCSFNISTIRGKQKEVAFYKLDIIPIDNTTISYTLTSCNTSVITQACP 211  
 QY 111 KISFEPIPIHYCAPAGAILIKNDKTFNGGPKCNVSTVOCGIRVYSTOLLNSLA 170  
 DB 212 KVSFEPIPIHYCAPAGAILIKNDKTFNGGPKCNVSTVOCGIRVYSTOLLNSLA 271  
 QY 171 EEEVVISDFTNNAKTIIVOLKESVEINCTRPNNTRKSIH--GPRAFYTTGEIIGD 228  
 DB 272 EEEVVISDFTNNAKTIIVOLKESVEINCTRPNNTRKSIH--GPRAFYTTGEIIGD 330  
 QY 229 IQAHNCNISRANKNDTLKQIVIKLRQEFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYC 287  
 DB 331 IQAHNCNISRANKNDTLKQIVIKLRQEFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYC 390  
 QY 288 NSTQLEFNSTNN--NTEGSNNTEG--NTITLPCRIOIIMMOEVEKAMVAPPIRQIR 343  
 DB 391 NSTQLEFNSTNN--NTEGSNNTEG--NTITLPCRIOIIMMOEVEKAMVAPPIRQIR 450  
 QY 344 SSNTTGLLTRDGGIENGTEIFRPGGDMRDWRSELYKVKVTEPIGVAPTKRRV 403  
 DB 451 SSNTTGLLTRDGGIENGTEIFRPGGDMRDWRSELYKVKVTEPIGVAPTKRRV 510  
 QY 404 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEM 463  
 DB 511 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEM 570  
 QY 464 LQITWGIKOLQARVLAVERYLDQOLLGIGCSGKLICTTAVPMAWSNKSLEDRIMN 523  
 DB 571 LQITWGIKOLQARVLAVERYLDQOLLGIGCSGKLICTTAVPMAWSNKSLEDRIMN 630  
 QY 524 MTWMEEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASIMNFDITNMLWY 579  
 DB 631 MTWMEEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASIMNFDITNMLWY 686

## RESULT 12

US-08-482-847-14  
 ; Sequence 14, Application US/08482847  
 ; Patent No. 5556757  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VAN ALSTYNE, Diane  
 ; APPLICANT: SHARMA, Lawrence Rajendra  
 ; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
 ; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
 ; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 40  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/482,847  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/127,499  
 ; FILING DATE: 28-SEP-1993  
 ; ATTORNEY/AGENT INFORMATION:

NAME: BERT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 51916/104/INBI  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 672-5300  
 TELEFAX: (202) 672-5399  
 TELER: 904136  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown

US-08-482-847-14

Query Match 87.1%; Score 2724.5; DB 1; Length 861;  
 Best Local Similarity 79.6%; Pred. No. 5.5e-227;  
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

2 EKLMTVVYGVVWKEATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLENTYE 61  
 32 EKLMTVVYGVVWKEATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLENTYE 91  
 62 HNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA-----GCDTSVITQACP 98  
 92 NNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA-----GCDTSVITQACP 151  
 99 -----GCDTSVITQACP 110  
 152 MMEKEIKNCSFNISTSIRKQVEAFYKLDIIPIDNDTTSYLTSCNTSVITQACP 211  
 111 KISFPIPIHYCAPAGFAILKCNKDTFNGKPCNKVSTVQCTHGIRPVVSTOLLNGSLA 170  
 212 KVSFPIPIHYCAPAGFAILKCNKDTFNGKPCNKVSTVQCTHGIRPVVSTOLLNGSLA 271  
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGAFYTTGEIIGD 228  
 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGAFYTTGEIIGD 330  
 229 IROAHCNISRAKMNLTQIIVIKLREOF--ENKTIYPNHSNGGDEIIVHMSFNGGGEFFYC 287  
 331 MROAHCNISRAKMNLTQIIVIKLREOF--ENKTIYPNHSNGGDEIIVHMSFNGGGEFFYC 390  
 288 NSTOLFNSTWNN---NTEGSNNTG--NTITLPCRIRKQIINMWOEVGKAMYPAPPIGOIRC 343  
 391 NSTOLFNSTWNN---NTEGSNNTG--NTITLPCRIRKQIINMWOEVGKAMYPAPPIGOIRC 450  
 344 SSNITGLLITRDGINENGTETIFRPGGDMRDNRWSELYKYKVXKIEPLGVAPTYCKRRV 403  
 451 SSNITGLLITRDGINENGTETIFRPGGDMRDNRWSELYKYKVXKIEPLGVAPTYCKRRV 510  
 404 VOREKRAVIGAVFLGFLGAAGSTWGAASMTLVQARLLSGIVQOONNLLRAIEQOQM 463  
 511 VOREKRAVIGAVFLGFLGAAGSTWGAASMTLVQARLLSGIVQOONNLLRAIEQOQM 570  
 464 LQLTWGIQOLARVLAVERYLGDQDLGIGGSGCLICTTAVPMWASNSKSLDRIMN 523  
 571 LQLTWGIQOLARVLAVERYLGDQDLGIGGSGCLICTTAVPMWASNSKSLDRIMN 630  
 524 MTMWEREIDNYTSEIYTLIEBSONQOEKNEOELLEDKMASLWMPDITWLMW 579  
 631 MTMWEREIDNYTSEIYTLIEBSONQOEKNEOELLEDKMASLWMPDITWLMW 686

RESULT 13  
 US-07-956-483-10  
 Sequence 10, Application US/07956483  
 Patent No. 6261799  
 GENERAL INFORMATION:  
 APPLICANT: KIENY, Marie-Paule  
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
 NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/956,483  
 FILING DATE: 31-DEC-1992  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO 92/19742  
 FILING DATE: 12-NOV-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 91 05392  
 FILING DATE: 02-MAY-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 01753-005  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-07-956-483-10

Query Match 87.1%; Score 2724.5; DB 3; Length 861;  
 Best Local Similarity 79.6%; Pred. No. 5.5e-227;  
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

2 EKLMTVVYGVVWKEATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLENTYE 61  
 32 EKLMTVVYGVVWKEATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPQEVLENTYE 91  
 62 HNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA-----GCDTSVITQACP 98  
 92 NNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA-----GCDTSVITQACP 151  
 99 -----GCDTSVITQACP 110  
 152 MMEKEIKNCSFNISTSIRKQVEAFYKLDIIPIDNDTTSYLTSCNTSVITQACP 211  
 111 KISFPIPIHYCAPAGFAILKCNKDTFNGKPCNKVSTVQCTHGIRPVVSTOLLNGSLA 170  
 212 KVSFPIPIHYCAPAGFAILKCNKDTFNGKPCNKVSTVQCTHGIRPVVSTOLLNGSLA 271  
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGAFYTTGEIIGD 228  
 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGAFYTTGEIIGD 330  
 229 IROAHCNISRAKMNLTQIIVIKLREOF--ENKTIYPNHSNGGDEIIVHMSFNGGGEFFYC 287  
 331 MROAHCNISRAKMNLTQIIVIKLREOF--ENKTIYPNHSNGGDEIIVHMSFNGGGEFFYC 390  
 288 NSTOLFNSTWNN---NTEGSNNTG--NTITLPCRIRKQIINMWOEVGKAMYPAPPIGOIRC 343  
 391 NSTOLFNSTWNN---NTEGSNNTG--NTITLPCRIRKQIINMWOEVGKAMYPAPPIGOIRC 450  
 344 SSNITGLLITRDGINENGTETIFRPGGDMRDNRWSELYKYKVXKIEPLGVAPTYCKRRV 403  
 451 SSNITGLLITRDGINENGTETIFRPGGDMRDNRWSELYKYKVXKIEPLGVAPTYCKRRV 510

Qy 404 VOREKRAVGIGAVELFGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQM 463  
Db 511 VOREKRAVGIGALFLGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOHL 570  
Qy 464 LQLTWVGIKQLOARVLAVERYLGDQOLLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 523  
Db 571 LQLTWVGIKQLOARVLAVERYLKDQOLLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 630  
Qy 524 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKMASLMMNFDTITNMLY 579  
Db 631 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKMASLMMNFDTITNMLY 686

## RESULT 14

US-08-472-240A-1

Sequence 1, Application US/08472240A

Patent No. 6284248

GENERAL INFORMATION:

APPLICANT: KIENY, Marie-Paule

TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,240A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,483

FILING DATE: 31-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-055

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 861 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Protein

LOCATION: 1..854

US-08-472-240A-1

Query Match 87.1%; Score 2724.5; DB 3; Length 861;

Best Local Similarity 79.6%; Pred. No. 5.5e-227;

Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

Qy 2 EKLWTVVYGVVPMKEATTLTFCASDAKAYDTEVHNHVMATHACVPTDPNPQSEVLEANTY 61  
Db 32 EKLWTVVYGVVPMKEATTLTFCASDAKAYDTEVHNHVMATHACVPTDPNPQSEVLEANTY 91  
Qy 62 HNNMKNMVEQMODIISLMQSLKPCVKLTPLGCA----- 98  
Db 92 NFNMKNDVEMQHEDIISLMQSLKPCVKLTPLGCA----- 151

Qy 99 -----GDTSVITQACP 110  
Db 152 MMKEGEIKNCSEFNISTSRGKVQKEYAFYKLDIIPIDNDTTSYLLSCNTSVITQACP 211  
Qy 111 KISFEPIPIHYCAPAGAILKQNDKPFNGPCKNYSTVOCCTGIRPVSTOLLNSLA 170  
Db 212 KVSFEPIPIHYCAPAGAILKCNKTFNGPCPNYSTVOCCTGIRPVSTOLLNSLA 271  
Qy 171 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH- -GPRAFYTTGELIGD 228  
Db 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH- -GPRAFYTTGELIGD 330  
Qy 229 IROAHNCISRAKNDTLKQIVIKREOF-ENKTIIVFNHSGGDEIYMSFNGGEFFYC 287  
Db 331 IROAHNCISRAKNDTLKQIVIKREOF-ENKTIIVFNHSGGDEIYMSFNGGEFFYC 390  
Qy 288 NSTOLENSTWNN--NTEGSSNTEG--NTITLPCRKQIIMMOEYGVAMAPRQGRV 343  
Db 391 NSTOLENSTWNN--NTEGSSNTEG--NTITLPCRKQIIMMOEYGVAMAPRQGRV 450  
Qy 344 SSNTGLLTPRDGINENGTEIFPPGGGDMRDMRSELYKVKYKLEPLGVAPTKGRV 403  
Db 451 SSNTGLLTPRDGINENGTEIFPPGGGDMRDMRSELYKVKYKLEPLGVAPTKGRV 510  
Qy 404 VOREKRAVGIGAVELFGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQM 463  
Db 511 VOREKRAVGIGALFLGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOHL 570  
Qy 464 LQLTWVGIKQLOARVLAVERYLGDQOLLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 523  
Db 571 LQLTWVGIKQLOARVLAVERYLKDQOLLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 630  
Qy 524 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKMASLMMNFDTITNMLY 579  
Db 631 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKMASLMMNFDTITNMLY 686

## RESULT 15

US-08-472-240A-7

Sequence 7, Application US/08472240A

Patent No. 6284248

GENERAL INFORMATION:

APPLICANT: KIENY, Marie-Paule

TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,240A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,483

FILING DATE: 31-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-055

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 861 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..831  
US-08-472-240A-7

Query Match

87.1%; Score 2724.5; DB 3; length 861;

Best Local Similarity 79.6%; Pred. No. 5,5e-227;

Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

QY 2 EKLMTVVYGVPPWKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVTE 61  
DB 32 EKLMTVVYGVPPWKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVTE 91  
62 HFMNMKNNVQWQEDIIISLMDQSLKPCVLTPLCGA----- 98  
92 NFMNMKNNDVNEQWHEHIIISLMDQSLKPCVLTPLCVSLKCTDLGNATNTNNTSSSGE 151  
99 -----GCDTSVITQACP 110  
DB 152 MMEKEIKNCSEFNISTIRGKYQKEVAFYKLDIIPINDTTSYLTSCNTSVITQACP 211  
QY 111 KISFEPPIHYCAPAGPAILKCKNDKTENGKPCKNVSTVOCTHGRPVVSTOLLNGSLA 170  
DB 212 KVSFEPPIHYCAPAGPAILKCKNNKTFNGTGVCTVSTVOCTHGRPVVSTOLLNGSLA 271  
QY 171 EEEVVRSDNFTNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTEIGD 228  
DB 272 EEEVVRSDNFTNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTEIGD 330  
QY 229 IROAHNCISRAKKNLTKQIVIKLRQF--ENKTIIVNHSQGDPEIWMHSFNGGEFFYC 287  
DB 331 MROAHNCISRAKKNLTKQIVIKLRQF--ENKTIIVNHSQGDPEIWMHSFNGGEFFYC 390  
QY 288 NSTQLFNSTWNN--NTEGSNNTEG--NTTLPCRIOIIMNMQEVGKAMYPPIRGQIRG 343  
DB 391 NSTQLFNSTWNN--NTEGSNNTEG--NTTLPCRIOIIMNMQEVGKAMYPPIRGQIRG 450  
QY 344 SSNTGLLTLTRDGINENGTEIFRPGGDMRDNRSELYKVKIIEPLGVAPTCKRRY 403  
DB 451 SSNTGLLTLTRDGINENGTEIFRPGGDMRDNRSELYKVKIIEPLGVAPTCKRRY 510  
QY 404 VOREKRAVGIGAVFLGFLGAAGSTMGASWTLTVQARLLLSGIVQOQNNLLRAIEAQRM 463  
DB 511 VOREKRAVGIGAVFLGFLGAAGSTMGASWTLTVQARLLLSGIVQOQNNLLRAIEAQRM 570  
QY 464 LQUTWGIKQOLQARVLAVERYLDDQOLLGIGCGSKLICCTAVPMNASWSNKSLEIRIANN 523  
DB 571 LQUTWGIKQOLQARVLAVERYLDDQOLLGIGCGSKLICCTAVPMNASWSNKSLEIRIANN 630  
QY 524 MTMMEWEREIDNVTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITNMLWY 579  
DB 631 MTMMEWEREIDNVTSEIYTLIEESQNOQEKNEOELLELDKMASLWMPDITNMLWY 686

Search completed: December 12, 2003, 12:34:28  
Job time : 15.1662 secs

**THIS PAGE BLANK (US.110)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 24.1381 Seconds  
(without alignments)  
4461.192 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129  
Sequence: 1 VEKLMVTYYGVVPWKKEATT.....ELDKWASIMWFDTITMLMY 579

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3129	100.0	579	US-10-032-162-15	Sequence 15, Appl
2	3056	97.7	643	US-10-032-162-13	Sequence 17, Appl
3	2923	93.4	625	US-10-032-162-17	Sequence 17, Appl
4	2771	88.6	619	US-09-891-609-4	Sequence 4, Appl
5	2771	88.6	646	US-09-891-609-2	Sequence 2, Appl
6	2764	88.3	842	US-10-190-435-2	Sequence 2, Appl
7	2764	88.3	842	US-10-241-009-2	Sequence 2, Appl
8	2764	88.3	842	US-10-190-434B-2	Sequence 2, Appl
9	2764	88.3	842	US-10-190-305A-2	Sequence 2, Appl
10	2764	88.3	847	US-09-476-242-2	Sequence 2, Appl
11	2724.5	87.1	861	US-10-026-741-103	Sequence 103, App
12	2702	86.4	856	US-09-476-242-1	Sequence 1, Appl
13	2694	86.1	856	US-10-196-515-11	Sequence 11, Appl
14	2672.5	85.4	726	US-10-196-515-3	Sequence 3, Appl
15	2651.5	84.7	759	US-10-196-515-12	Sequence 12, Appl

16	2641.5	84.4	853	12	US-10-286-332A-33	Sequence 33, Appl
17	2641.5	84.4	853	12	US-10-280-915-33	Sequence 33, Appl
18	2641.5	84.4	853	14	US-10-003-035-33	Sequence 33, Appl
19	2625.5	83.9	868	10	US-09-938-406-1	Sequence 1, Appl
20	2535	81.0	1101	12	US-10-286-332A-53	Sequence 53, Appl
21	2535	81.0	1101	12	US-10-280-915-53	Sequence 53, Appl
22	2535	81.0	1101	14	US-10-003-035-53	Sequence 53, Appl
23	2520	80.5	1186	12	US-10-286-332A-55	Sequence 55, Appl
24	2520	80.5	1186	12	US-10-280-915-55	Sequence 55, Appl
25	2520	80.5	1186	14	US-10-003-035-55	Sequence 55, Appl
26	2495.5	79.8	860	12	US-10-190-435-6	Sequence 6, Appl
27	2495.5	79.8	860	12	US-10-241-009-6	Sequence 6, Appl
28	2495.5	79.8	860	12	US-10-190-434B-6	Sequence 6, Appl
29	2495.5	79.8	860	12	US-10-190-305A-6	Sequence 6, Appl
30	2486	79.5	858	12	US-10-190-435-150	Sequence 150, App
31	2457	78.5	867	12	US-10-190-435-3	Sequence 3, Appl
32	2457	78.5	867	12	US-10-190-435-126	Sequence 126, App
33	2457	78.5	867	12	US-10-241-009-3	Sequence 3, Appl
34	2457	78.5	867	12	US-10-190-434B-3	Sequence 3, Appl
35	2457	78.5	867	12	US-10-190-305A-3	Sequence 3, Appl
36	2449	78.3	870	12	US-10-190-435-147	Sequence 147, App
37	2448.5	78.3	803	12	US-10-190-435-135	Sequence 135, App
38	2447.5	78.2	803	12	US-10-190-435-134	Sequence 134, App
39	2447	78.2	861	12	US-10-190-435-139	Sequence 139, App
40	2446	78.2	844	10	US-09-991-258-19	Sequence 19, Appl
41	2442.5	78.1	845	12	US-10-190-435-129	Sequence 129, App
42	2442.5	78.1	845	12	US-10-190-435-130	Sequence 130, App
43	2440.5	78.0	845	12	US-10-190-435-143	Sequence 143, App
44	2439	77.9	869	12	US-10-190-435-4	Sequence 4, Appl
45	2439	77.9	869	12	US-10-241-009-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-10-032-162-15  
; Sequence 15, Application US/10032162  
; Publication No US20030052839A1  
GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADON J  
; APPLICANT: JOHN, MOORE P  
TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
FILE REFERENCE: 2048/5931a2  
CURRENT APPLICATION NUMBER: US/10/032.162  
PRIOR FILING DATE: 2002-10-24  
PRIOR APPLICATION NUMBER: 09/602,864  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-15  
Query Match 100.0%; Score 3129; DB 15; Length 579;  
Best Local Similarity 100.0%; Pred. No. 6.9e-287;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VEKLMVTYYGVVPWKKEATTLTFCASDAKAYDEVHNWATHACVPTDPNQEVLLENT 60  
DB 1 VEKLMVTYYGVVPWKKEATTLTFCASDAKAYDEVHNWATHACVPTDPNQEVLLENT 60  
QY EHFMKKNNWQKQEDISLMDOSLKPVCVLTPLCGAGCOTSVITOCAPKISPEPIYH 120  
DB EHFMKKNNWQKQEDISLMDOSLKPVCVLTPLCGAGCOTSVITOCAPKISPEPIYH 120  
QY 121 YCAGAGFALCNKTKTPGKGPCKNVSIVQCTHGIRPVSTOLLNGSLAEVEVIRSDN 180  
DB 121 YCAGAGFALCNKTKTPGKGPCKNVSIVQCTHGIRPVSTOLLNGSLAEVEVIRSDN 180

Db 121 YCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVIRSDN 180  
Qy 181 FTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRAHONISRAK 240  
Db 181 FTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRAHONISRAK 240  
Qy 241 WNDLTKQIVIKLREQENKTIIVFNHSSGGDEPEIVHNSFNGGEFFYCNSITOLFNSITMNN 300  
Db 241 WNDLTKQIVIKLREQENKTIIVFNHSSGGDEPEIVHNSFNGGEFFYCNSITOLFNSITMNN 300  
Qy 301 TEGSNNTEGNTITLPCRKOIINMVOEYKAMVAPPIRGQIRCSSNITGLLTRDGGINE 360  
Db 301 TEGSNNTEGNTITLPCRKOIINMVOEYKAMVAPPIRGQIRCSSNITGLLTRDGGINE 360  
Qy 361 NGTEIFRPGGDMRSELYKVKYKIEPLGVAFTKCKRRVQREKRAVIGAVLGF 420  
Db 361 NGTEIFRPGGDMRSELYKVKYKIEPLGVAFTKCKRRVQREKRAVIGAVLGF 420  
Qy 421 LGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTVWGIKQLOA 480  
Db 421 LGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTVWGIKQLOA 480  
Qy 481 VERYLGDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWMEEREIDNYTSEI 540  
Db 481 VERYLGDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWMEEREIDNYTSEI 540  
Qy 541 YTLIESONOEKNEOELLELDKASLMMNPDITNMLWY 579  
Db 541 YTLIESONOEKNEOELLELDKASLMMNPDITNMLWY 579

## RESULT 2

US-10-032-162-13  
; Sequence 13, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUEBKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331az  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 643  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-032-162-13

Query Match 97.7%; Score 3056; DB 15; Length 643;  
Best Local Similarity 89.3%; Pred. No. 6.5e-280;  
Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;  
Qy 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60  
Db 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60  
Qy 61 EHFNMKNKNWTEQMEDIIISLMDOSLKPCVKLTPLCVLJNCKDVNATNTNDSGTMERG 98  
Db 61 EHFNMKNKNWTEQMEDIIISLMDOSLKPCVKLTPLCVLJNCKDVNATNTNDSGTMERG 120  
Qy 99 -----GCDTSVITACPKISFEP 116  
Db 121 EIKKCSFNTTISIRDEVQKVALFKLDVVPIDNNNTSYRLISCTSVITQACPKISFEP 180  
Qy 117 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVIR 176  
Db 117 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVIR 176

Db 181 IPIHYCAPAGFAILKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVIR 240  
Qy 177 RSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRAHONISRAK 236  
Db 241 RSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRAHONISRAK 300  
Qy 237 SRAKMDTLKQIVIKLREQENKTIIVFNHSSGGDEPEIVHNSFNGGEFFYCNSITOLFNSIT 236  
Db 301 SRAKMDTLKQIVIKLREQENKTIIVFNHSSGGDEPEIVHNSFNGGEFFYCNSITOLFNSIT 360  
Qy 297 WNNTEGSNNTEGNTITLPCRKOIINMVOEYKAMVAPPIRGQIRCSSNITGLLTRDGGINE 356  
Db 361 WNNTEGSNNTEGNTITLPCRKOIINMVOEYKAMVAPPIRGQIRCSSNITGLLTRDGGINE 420  
Qy 357 GINENGTEIFRPGGDMRSELYKVKYKIEPLGVAFTKCKRRVQREKRAVIGAVLGF 416  
Db 421 GINENGTEIFRPGGDMRSELYKVKYKIEPLGVAFTKCKRRVQREKRAVIGAVLGF 480  
Qy 417 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTVWGIKQLOA 476  
Db 481 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTVWGIKQLOA 540  
Qy 477 RVLAVERYLGDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWMEEREIDNYTSEI 536  
Db 541 RVLAVERYLGDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWMEEREIDNYTSEI 600  
Qy 537 TSELYTLIESONOEKNEOELLELDKASLMMNPDITNMLWY 579  
Db 601 TSELYTLIESONOEKNEOELLELDKASLMMNPDITNMLWY 643

## RESULT 3

US-10-032-162-17  
; Sequence 17, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUEBKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADDON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331az  
; CURRENT APPLICATION NUMBER: US/10/032,162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 625  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (151)..(151)  
; OTHER INFORMATION: X=UNKNOWN AMINO ACID  
; NAME/KEY: MISC FEATURE  
; LOCATION: (205)..(205)  
; OTHER INFORMATION: X=UNKNOWN AMINO ACID  
; NAME/KEY: MISC FEATURE  
; LOCATION: (213)..(213)  
; OTHER INFORMATION: X=UNKNOWN AMINO ACID  
US-10-032-162-17

Query Match 93.4%; Score 2923; DB 15; Length 625;  
Best Local Similarity 86.0%; Pred. No. 2.3e-267;  
Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;  
Qy 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60  
Db 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60



```

Db      1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
Qy      61 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCGA----- 98
Db      61 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCVTLNCKDVNATNTTNDSEGMERG 120
Qy      99 -----GCDTSVITQACPISFEP 116
Db      121 EIKNCSFNITTSIRDEVQKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPISFEP 180
Qy      117 IPIHCAPGFAILKCNKDTFNGKPCRVSTVQCTHGRIPVSVTOLLNGLAEBEYV 176
Db      181 IPIHCAPGFAILKCNKDTFNGKPCRVSTVQCTHGRIPVSVTOLLNGLAEBEYV 240
Qy      177 RSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGGRAFYTTGELIGIROAHCHI 236
Db      241 RSDNFTNNAKTIIVOLKESVEINCRPNNTKSHIGGRAFYTTGELIGIROAHCHI 282
Qy      237 SRAKNDTLKOIIVIKLREQFENKTIIVFNHSGGDEBEIVMHSFNCGEFPYCNSDTLPNST 296
Db      283 SRAKNDTLKOIIVIKLREQFENKTIIVFNHSGGDEBEIVMHSFNCGEFPYCNSDTLPNST 342
Qy      297 WNNNTEGNTNTEGNTITLPCRIKOIIMMOEYKAMYPPIRGQIRCSNITGLLTRDG 356
Db      343 WNNNTEGNTNTEGNTITLPCRIKOIIMMOEYKAMYPPIRGQIRCSNITGLLTRDG 402
Qy      357 GINENGTEIFRPGGDMRDMRSELYKYKVYKIEPIGAVPTCKRVRVOREKRAVGIGAV 416
Db      403 GINENGTEIFRPGGDMRDMRSELYKYKVYKIEPIGAVPTCKRVRVOREKRAVGIGAV 462
Qy      417 FLGFLGAAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQORMLQITVWGIKOLA 476
Db      463 FLGFLGAAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQORMLQITVWGIKOLA 522
Qy      477 RVLAVERYLGDOQLLGIWCGSGKLCCTRAVPNMASSNKSILDRINNNMTWMEWEIEDNY 536
Db      523 RVLAVERYLGDOQLLGIWCGSGKLCCTRAVPNMASSNKSILDRINNNMTWMEWEIEDNY 582
Qy      537 TSEIYTLIEESONQOEKNEOBELLEDKMASLNMWFDITNWLMTY 579
Db      583 TSEIYTLIEESONQOEKNEOBELLEDKMASLNMWFDITNWLMTY 625

RESULT 4
US-09-891-609-4
; Sequence 4, Application US/09891609
; Patent No. US20020127238A1
; GENERAL INFORMATION:
; APPLICANT: Stamatatos, Leonidas
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,608
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-4

Query Match      88.6%; Score 2771; DB 10; Length 619;
Best Local Similarity 85.6%; Pred. No. 5,5e-253;
Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

```

```

Qy      1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
Db      3 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 62
Qy      61 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLC----- 96

```

```

Db      63 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCVTLHCTNLKNAATNTKSSNMKEMDR 122
Qy      97 -----GAG-----CDTSVITQACPISFEPPIHYCAPAGFAILKCNKDTFNGKPC 142
Db      123 GEIKNCSRVGAGKILNCNTSVITQACRVSEFPIPIHYCAPAGFAILKCNKDTFNGKPC 182
Qy      143 CKNVSTVQCTHGRIPVSVTOLLNGLAEBEYVIRSDNFTNNAKTIIVOLKESVEINCR 202
Db      183 CKNVSTVQCTHGRIPVSVTOLLNGLAEBEYVIRSEFTDPAKTIIVOLKESVEINCR 242
Qy      203 PNNNTRKSHIGGRAFYTTGELIGIROAHCHNISRAKNDTLKOIIVIKLREQFENKTIIV 262
Db      243 PNNNTRKSHIGGRAFYTTGELIGIROAHCHNISGEKNNNTLKOIIVIKLREQFENKTIIV 302
Qy      263 FHNHSGGDEBEIVMHSFNCGEFPYCNSDTLPNSTNNNTNTEGNTITLPCRIKOI 322
Db      303 FHNHSGGDEBEIVMHSFNCGEFPYCNSDTLPNSTNNNTNTEGNTITLPCRIKOI 360
Qy      323 NMMOEVGKAMYPPIRGQIRCSNITGLLTRDGGIN-ENGTEIFRPGGDMRDMRSEL 381
Db      361 NMMOEVGKAMYPPIRGQIRCSNITGLLTRDGKEISNTTEIFRPGGDMRDMRSEL 420
Qy      382 YKYKVYKIEPIGAVPTCKRVRVOREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARL 441
Db      421 YKYKVYKIEPIGAVPTCKRVRVOREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARL 480
Qy      442 LLSGIVQOONNLRAIEAQORMLQITVWGIKOLARVLAVERYLGDQOLLGIWCGSGKLI 501
Db      481 LLSGIVQOONNLRAIEAQORMLQITVWGIKOLARVLAVERYLGDQOLLGIWCGSGKLI 539
Qy      502 CCTRAVPNMASSNKSILDRINNNMTWMEWEIEDNTSEIYTLIEESONQOEKNEOBELLED 561
Db      540 CCTRAVPNMASSNKSILDRINNNMTWMEWEIEDNTSEIYTLIEESONQOEKNEOBELLED 599
Qy      562 DKMASLNMWFDITNWLMTY 579
Db      600 DKMASLNMWFDITNWLMTY 617

RESULT 5
US-09-891-609-2
; Sequence 2, Application US/09891609
; Patent No. US20020127238A1
; GENERAL INFORMATION:
; APPLICANT: Stamatatos, Leonidas
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,608
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-2

Query Match      88.6%; Score 2771; DB 10; Length 646;
Best Local Similarity 85.6%; Pred. No. 5,8e-253;
Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

```

```

Qy      1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
Db      30 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 89
Qy      61 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLC----- 96
Db      90 EHFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCTTLHCTNLKNAATNTKSSNMKEMDR 149
Qy      97 -----GAG-----CDTSVITQACPISFEPPIHYCAPAGFAILKCNKDTFNGKPC 142

```

```

Db 150 GEIKNCSEFKVAGAKLINCNTSVITQACPVSFEPIPIHYCAPAGFALIKCNCKKXNGSGP 209
Qy 143 CKNVSTQCTHGIRPVYSTOLLNGSLAESEVIRSDNFTNNAKTIIVQLSEVEINSTR 202
Db 210 CTNVSTQCTHGIRPVYSTOLLNGSLAESEVIRSENFTNNAKTIIVQLSEVEINSTR 269
Qy 203 PNNTRKSIHIGPRAFYTTGEIIGDIRQAHCHNISRANKNDLKOIVIKLEOPENKTIV 262
Db 270 PNNTRKSIITGPRAFYATGDIIGDIRQAHCHNISGEKMNNTLKOIVIKLEOPENKTIV 329
Qy 263 FNHSSGDPPEIVHMSFNGCGEFPYCNSTQTFNSNTMANNTEGNTITLPCRIKOI 322
Db 330 FKOSGDPPEIVHMSFNGCGEFPYCNSTQTFNSNTMANNTEGNTITLPCRIKOI 387
Qy 323 NNMGEVGAAPPIRGOIRCSSNTGLLTRDGIN-ENGTEIFRPGGDMRDNRSEL 381
Db 388 NNMGEVGAAPPIRGOIRCSSNTGLLTRDGIN-ENGTEIFRPGGDMRDNRSEL 447
Qy 382 YKYVVKIEPLGVAPTKCRKRVOREKRAVGIGAVFLGFGAAGSTWGAASMTLTVQARL 441
Db 448 YKYVVKIEPLGVAPTKCRKRVOREKRAVGIGAVFLGFGAAGSTWGAASMTLTVQARL 507
Qy 442 LLSGIVQOQNNLLRAIEAQOQMLQTLTWGIRKQOLAVERYIGDQQLGIGWCSGKLI 501
Db 508 LLSGIVQOQNNLLRAIEAQOQMLQTLTWGIRKQOLAVERYIGDQQLGIGWCSGKLI 566
Qy 502 CCTAVPNNASWSKSLDRIANNMTMWEEREIDNTSEIYTLIEESONQOEKNEOELEL 561
Db 567 CCTAVPNNASWSKSLDRIANNMTMWEEREIDNTSEIYTLIEESONQOEKNEOELEL 626
Qy 562 DKMASLNMWPDITNMLWY 579
Db 627 DKMASLNMWPDITNMLWY 644

```

```

RESULT 6
US-10-190-435-2
; Sequence 2, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-190-435-2

```

```

Query Match 88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy 1 VEKLMVTYVYGVVPWKAKATITLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLENTV 60
Db 25 VEKLMVTYVYGVVPWKAKATITLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLENTV 84
Qy 61 EHFNMKNMNVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 98
Db 85 EHFNMKNMNVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 144
Qy 99 -----GCDTSVITQACPKISFE 115

```

```

Db 145 GEIKNCSEFKVTSIRNMQKEVALFYKLDVPIPDNDNTSYKLIINCNTSVITQACPVSFE 204
Qy 116 PIPHYCAPAGFALIKCNCKKXNGSGPCKRVSTVQCTHGIRPVYSTOLLNGSLAESEVIR 175
Db 205 PIPHYCAPAGFALIKCNCKKXNGSGPCKRVSTVQCTHGIRPVYSTOLLNGSLAESEVIR 264
Qy 176 IRSNFTNNAKTIIVQLSEVEINSTRPNNTRKSIHIGPRAFYTTGEIIGDIRQAHCHN 235
Db 265 IRSNFTNNAKTIIVQLSEVEINSTRPNNTRKSIITGPRAFYATGDIIGDIRQAHCHN 324
Qy 236 ISRAKNDTLKOIVIKLEOPENKTIVFNHSSGDPPEIVHMSFNGCGEFPYCNSTQTFNS 295
Db 325 ISRAKNDTLKOIVIKLEOPENKTIVFNHSSGDPPEIVHMSFNGCGEFPYCNSTQTFNS 384
Qy 296 TMMNTEGNTITLPCRIKOIINNMGEVGAAPPIRGOIRCSSNTGLLTRD 355
Db 385 TW-NNTIGPNNNTNG-TITLPCRIKOIINNMGEVGAAPPIRGOIRCSSNTGLLTRD 442
Qy 356 GGIN-ENGTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTKCRKRVVQREKRAVGIG 414
Db 443 GKEISNTTEIFRPGGDMRDNRSELKYKVKVIEPLGVAPTKCRKRVVQREKRAVGIG 502
Qy 415 AVFLGFGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQTLTWGIRKOL 474
Db 503 AVFLGFGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQTLTWGIRKOL 562
Qy 475 QARVLAVERYIGDQQLGIGWCSGKLICTTAVPNNASWSKSLDRIANNMTMWEEREID 534
Db 563 QARVLAVERYIGDQQLGIGWCSGKLICTTAVPNNASWSKSLDRIANNMTMWEEREID 622
Qy 535 NYTSEIYTLIEESONQOEKNEOELELBDKMASLNMWPDITNMLWY 579
Db 623 NYTSEIYTLIEESONQOEKNEOELELBDKMASLNMWPDITNMLWY 667

```

```

RESULT 7
US-10-241-009-2
; Sequence 2, Application US/10241009
; Publication No. US20030170614A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.21
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-241-009-2

```

```

Query Match 88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy 1 VEKLMVTYVYGVVPWKAKATITLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLENTV 60
Db 25 VEKLMVTYVYGVVPWKAKATITLFCASDAKAYDTEVHNWATHACVPTDPNQEIVLENTV 84
Qy 61 EHFNMKNMNVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 98
Db 85 EHFNMKNMNVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 144
Qy 99 -----GCDTSVITQACPKISFE 115

```

```

Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204
Qy      116 PIPHYCAPAGFAILKNDKTEGNGKPGKGNVSTVOCTHGIRPVSTOLLNLSLAEEVY 175
Db      205 PIPHYCAPAGFAILKNDKTEGNGKPGKGNVSTVOCTHGIRPVSTOLLNLSLAEEVY 264
Qy      176 IRSDFNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEEIIGDIRQAHCN 235
Db      265 IRSEFTDPAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAHCN 324
Qy      226 ISRAKNDTLKOIIVKLAEQFENKTIIVFNHSSGGDEPEIYMSFNGGGEFFYCNSTOLFNS 295
Db      325 ISGEKMNNTLKOIIVKLAEQFENKTIIVFQSSGGDEPEIYMSFNGGGEFFYCNSTOLFNS 384
Qy      226 TNNNTGSGNNTEGNTITLPCRIKOIIMMOEVRGAMVAPPIRGQIRCSNTITGLLTFRD 355
Db      385 TW-ANTIGENNTNG-TITLPCRIKOIIMMOEVRGAMVAPPIRGQIRCSNTITGLLTFRD 442
Qy      356 GGIN-ENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAFTKCRRVQREKRAVGIG 414
Db      443 GKEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAFTKCRRVQREKRAVTLG 502
Qy      415 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLDLTWGIGIKOL 474
Db      503 AMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNNLLRAIEAQOHLDLTWGIGIKOL 562
Qy      475 QARVLAVERYLGDQDLGIGWCSGKLICTTAVPNNASNSKSLDRIIMNMTMWEWERID 534
Db      563 QARVLAVERYLKDQDLGIGWCSGKLICTTAVPNNASNSKSLDRIIMNMTMWEWERID 622
Qy      535 NYTSEIYTLIESQNOQEKNEQELLEDKMASLNMWFDITNMLWY 579
Db      623 NYTNLIYTLIESQNOQEKNEQELLEDKMASLNMWFDISKMLWY 667

```

```

RESULT 8
US-10-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARRETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/10/190,434B
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SP162
US-10-190-434B-2

```

```

Query Match      88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy      1 VEKLVVTVYGVVWKEATTTLPFCASDAKAYDTEVHNWATACVPTDNPQEVYLENT 60
Db      25 VEKLVVTVYGVVWKEATTTLPFCASDAKAYDTEVHNWATACVPTDNPQEVYLENT 84
Qy      61 EHFNNKNNMVEQOMEDIISLMDQSLKPCVKLTPLCGA----- 98
Db      85 EHFNNKNNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLHCTNLKNAITNTSSNNKEMDR 144
Qy      99 -----GCDTSVITQACPKVSFE 115
Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204

```

```

Qy      116 PIPHYCAPAGFAILKNDKTEGNGKPGKGNVSTVOCTHGIRPVSTOLLNLSLAEEVY 175
Db      205 PIPHYCAPAGFAILKNDKTEGNGKPGKGNVSTVOCTHGIRPVSTOLLNLSLAEEVY 264
Qy      176 IRSDFNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEEIIGDIRQAHCN 235
Db      265 IRSEFTDPAKTIIVQLKESVEINCTRPNNNTRKSIITIGPRAFYATGDIIGDIRQAHCN 324
Qy      226 ISRAKNDTLKOIIVKLAEQFENKTIIVFNHSSGGDEPEIYMSFNGGGEFFYCNSTOLFNS 295
Db      325 ISGEKMNNTLKOIIVKLAEQFENKTIIVFQSSGGDEPEIYMSFNGGGEFFYCNSTOLFNS 384
Qy      226 TNNNTGSGNNTEGNTITLPCRIKOIIMMOEVRGAMVAPPIRGQIRCSNTITGLLTFRD 355
Db      385 TW-ANTIGENNTNG-TITLPCRIKOIIMMOEVRGAMVAPPIRGQIRCSNTITGLLTFRD 442
Qy      356 GGIN-ENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAFTKCRRVQREKRAVGIG 414
Db      443 GKEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAFTKCRRVQREKRAVTLG 502
Qy      415 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLDLTWGIGIKOL 474
Db      503 AMFLGFLGAAGSTMGARSLTLTVQARQLLSGIVQOQNNLLRAIEAQOHLDLTWGIGIKOL 562
Qy      475 QARVLAVERYLGDQDLGIGWCSGKLICTTAVPNNASNSKSLDRIIMNMTMWEWERID 534
Db      563 QARVLAVERYLKDQDLGIGWCSGKLICTTAVPNNASNSKSLDRIIMNMTMWEWERID 622
Qy      535 NYTSEIYTLIESQNOQEKNEQELLEDKMASLNMWFDITNMLWY 579
Db      623 NYTNLIYTLIESQNOQEKNEQELLEDKMASLNMWFDISKMLWY 667

```

```

RESULT 9
US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARRETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SP162
US-10-190-305A-2

```

```

Query Match      88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy      1 VEKLVVTVYGVVWKEATTTLPFCASDAKAYDTEVHNWATACVPTDNPQEVYLENT 60
Db      25 VEKLVVTVYGVVWKEATTTLPFCASDAKAYDTEVHNWATACVPTDNPQEVYLENT 84
Qy      61 EHFNNKNNMVEQOMEDIISLMDQSLKPCVKLTPLCGA----- 98
Db      85 EHFNNKNNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLHCTNLKNAITNTSSNNKEMDR 144
Qy      99 -----GCDTSVITQACPKVSFE 115
Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204

```

QY 116 PIPHYCAPAGFALLKCNKDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEV 175  
 DB 205 PIPHYCAPAGFALLKCNKDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEV 264  
 QY 176 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFYATGDIIGDIRAHCN 235  
 DB 265 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFYATGDIIGDIRAHCN 324  
 QY 236 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTLPNS 295  
 DB 325 ISGKMNNTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTLPNS 384  
 QY 296 TANNTEGSNNTEGNTITLPCRIKQIINMOEYKAMYPPIRQIICSSNITGLLTRD 355  
 DB 385 TW-NNTIGPNNTNG-TITLPCRIKQIINMOEYKAMYPPIRQIICSSNITGLLTRD 442  
 QY 356 GGIN-ENGTEIFRGGGDMDNMRSELKYKVVIEPLGVAFTKRRVQREKRAVTLG 414  
 DB 443 GGEISNTTEIFRGGGDMDNMRSELKYKVVIEPLGVAFTKRRVQREKRAVTLG 502  
 QY 415 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRLTLTWGIKOL 474  
 DB 503 AMFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRLTLTWGIKOL 562  
 QY 475 QARVLAVERYLGDQDLGIMGCGSKLICCTAVPNNASMSNKSIDRINNNTMWMEREID 534  
 DB 563 QARVLAVERYLGDQDLGIMGCGSKLICCTAVPNNASMSNKSIDRINNNTMWMEREID 622  
 QY 535 NYTSEIYTLIEESONQOEKNEQELLEDKXASLWNPEDITNWLWY 579  
 DB 623 NYTSEIYTLIEESONQOEKNEQELLEDKXASLWNPEDITNWLWY 667

RESULT 10  
 US-09-476-242-2  
 ; Sequence 2, Application US/09476242  
 ; Patent No. US20020146683A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BARNETT, Susan  
 ; APPLICANT: HARTOG, Karin  
 ; APPLICANT: MARTIN, Eric  
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
 ; FILE REFERENCE: 1605.002  
 ; CURRENT APPLICATION NUMBER: US/09/476.242  
 ; CURRENT FILING DATE: 1999-12-30  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 847  
 ; TYPE: PRT  
 ; ORGANISM: Human immunodeficiency virus  
 US-09-476-242-2

Query Match 88.3%; Score 2764; DB 10; Length 847;  
 Best Local Similarity 81.7%; Pred. No. 4e-252;  
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;  
 QY 1 VEXIMVTVYGVVPWKATTTLLFCASDCAKADTEVHNWATHACVPDNPQOEVLNVT 60  
 DB 30 VEXIMVTVYGVVPWKATTTLLFCASDCAKADTEVHNWATHACVPDNPQOEVLNVT 89  
 QY 61 EHPNNNNNNNNWQOEIIISLMDOSLKPCKVLTPLGCA----- 98  
 DB 90 ENFMNNNNNNWQOEIIISLMDOSLKPCKVLTPLGCA----- 149  
 QY 99 -----GDDTSVITQACPKISFE 115  
 DB 150 GEIKNCSFKVTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKISFE 209  
 QY 116 PIPHYCAPAGFALLKCNKDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEV 175  
 DB 210 PIPHYCAPAGFALLKCNKDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEV 269

QY 176 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFYATGDIIGDIRAHCN 235  
 DB 270 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFYATGDIIGDIRAHCN 329  
 QY 236 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTLPNS 295  
 DB 330 ISGKMNNTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTLPNS 389  
 QY 296 TANNTEGSNNTEGNTITLPCRIKQIINMOEYKAMYPPIRQIICSSNITGLLTRD 355  
 DB 390 TW-NNTIGPNNTNG-TITLPCRIKQIINMOEYKAMYPPIRQIICSSNITGLLTRD 447  
 QY 356 GGIN-ENGTEIFRGGGDMDNMRSELKYKVVIEPLGVAFTKRRVQREKRAVTLG 414  
 DB 448 GGEISNTTEIFRGGGDMDNMRSELKYKVVIEPLGVAFTKRRVQREKRAVTLG 507  
 QY 415 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRLTLTWGIKOL 474  
 DB 508 AMFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRLTLTWGIKOL 567  
 QY 475 QARVLAVERYLGDQDLGIMGCGSKLICCTAVPNNASMSNKSIDRINNNTMWMEREID 534  
 DB 568 QARVLAVERYLGDQDLGIMGCGSKLICCTAVPNNASMSNKSIDRINNNTMWMEREID 627  
 QY 535 NYTSEIYTLIEESONQOEKNEQELLEDKXASLWNPEDITNWLWY 579  
 DB 628 NYTSEIYTLIEESONQOEKNEQELLEDKXASLWNPEDITNWLWY 672

RESULT 11  
 US-10-026-741-103  
 ; Sequence 103, Application US/10026741  
 ; Publication No. US20030049604A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHARNEAU, PIERRE  
 ; CLAVEL, FRANCOISE  
 ; BORMAN, ANDREW  
 ; OUILLENT, CAROLINE  
 ; GUETARD, DENISE  
 ; MONTAGNER, LUC  
 ; DONON DE SAINT-MARTIN, JACQUELINE  
 ; COHEN, JACQUES  
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR  
 ; SUBTYPE) ANTIGENS  
 ; NUMBER OF SEQUENCES: 103  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Finmeagan, Henderson, Farabow, Garrett &  
 ; Dunner, L.L.P.  
 ; STREET: 1300 I Street, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/026.741  
 ; FILING DATE: 27-Dec-2001  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/817,441  
 ; FILING DATE: 31-AUG-1998  
 ; APPLICATION NUMBER: PCT/FR 95/01391  
 ; FILING DATE: 20-OCT-1995  
 ; APPLICATION NUMBER: FR 9412554  
 ; FILING DATE: 20-OCT-1994  
 ; APPLICATION NUMBER: FR 9502526  
 ; FILING DATE: 03-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Meyers, Kenneth J.  
 ; REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 03260.6005-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4400  
 INFORMATION FOR SEQ ID NO: 103:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
 US-10-026-741-103

Query Match 87.1%; Score 2724.5; DB 15; Length 861;  
 Best Local Similarity 79.6%; Pred. No. 2,2e-248;  
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

QY	2	EKLAVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE	61
DB	32	EKLAVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE	91
QY	62	HNMMKNNMVEQMODIISLMDQSLKPCVKLTPLCGA-----	98
DB	92	NFMKNDMVEQMHEDIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTSSSGE	151
QY	99	-----GCDTSVITQACP	110
DB	152	MMMEGEIKGCSFNISTSRKVQKEYAFYKLDIIPIDNDTTSYKLTSCNTSVITQACP	211
QY	111	KISFEPIPIHYCAPAGFALIKCNDKTFNGKGPCKAVSTVQCTHGIRPVVSTOLLNLSIA	170
DB	212	KVSFPIPIHYCAPAGFALIKCNDKTFNGKGPCKAVSTVQCTHGIRPVVSTOLLNLSIA	271
QY	171	EEVVRIRSDNFNNMKTIIIVQLKESVEINCTRPNNNTKRSIH--GGRAFTTGEIIGD	228
DB	272	EEVVRIRSDNFNNMKTIIIVQLKESVEINCTRPNNNTKRSIH--GGRAFTTGEIIGD	330
QY	229	IRQAHCNISRAKANDTLKOIVIKLRQF--ENKTIFFNHSSGGDPEIYVMSFNGCGEFPYC	287
DB	331	IRQAHCNISRAKANDTLKOIVIKLRQF--ENKTIFFNHSSGGDPEIYVMSFNGCGEFPYC	390
QY	288	NSTQLFNSNTWNN--NTEGSSNTEG--NTITLPCRIOIIMMVEGKAMAPRISQIRIC	343
DB	391	NSTQLFNSNTWNN--NTEGSSNTEG--NTITLPCRIOIIMMVEGKAMAPRISQIRIC	450
QY	344	SSNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRV	403
DB	451	SSNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRV	510
QY	404	VOREGRAVIGAVFLGFLGAGSTWGAASMTLTVOARLLLSGIYVOOQNNLLRAIEAQRM	463
DB	511	VOREGRAVIGAVFLGFLGAGSTWGAASMTLTVOARLLLSGIYVOOQNNLLRAIEAQRM	570
QY	464	LQLTWGIKQLOARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASNSKSLDRIMNN	523
DB	571	LQLTWGIKQLOARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASNSKSLDRIMNN	630
QY	524	MTMWEEREIDNYSEIYTLIESONQOEKNEBELLELDKMASLMMWPDITNMLWY	579
DB	631	MTMWEEREIDNYSEIYTLIESONQOEKNEBELLELDKMASLMMWPDITNMLWY	686

RESULT 12  
 US-09-476-242-1  
 Sequence 1, Application US/09476242  
 Patent No. US2002014663A1  
 GENERAL INFORMATION:  
 APPLICANT: BARNETT, Susan  
 APPLICANT: HARTOG, Karin  
 APPLICANT: MARTIN, Eric  
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
 FILE REFERENCE: 1605.002

CURRENT APPLICATION NUMBER: US/09/476,242  
 CURRENT FILING DATE: 1999-12-30  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 1  
 LENGTH: 856  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus  
 US-09-476-242-1

Query Match 86.4%; Score 2702; DB 10; Length 856;  
 Best Local Similarity 79.6%; Pred. No. 3e-246;  
 Matches 518; Conservative 28; Mismatches 31; Indels 74; Gaps 6;

QY	2	EKLAVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE	61
DB	32	EKLAVTVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE	91
QY	62	HNMMKNNMVEQMODIISLMDQSLKPCVKLTPLCGA-----	98
DB	92	NFMKNDMVEQMHEDIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTSSSGE	151
QY	99	-----GCDTSVITQACP	115
DB	152	GEIKGCSFNISTSRKVQKEYAFYKLDIIPIDNDTTSYKLTSCNTSVITQACPVSPE	211
QY	116	PIPIHYCAPAGFALIKCNDKTFNGKGPCKAVSTVQCTHGIRPVVSTOLLNLSIAEEV	175
DB	212	PIPIHYCAPAGFALIKCNDKTFNGKGPCKAVSTVQCTHGIRPVVSTOLLNLSIAEEV	271
QY	176	IRSDNFNNMKTIIIVQLKESVEINCTRPNNNTKRSIH--GGRAFTTGEIIGD	233
DB	272	IRSDNFNNMKTIIIVQLKESVEINCTRPNNNTKRSIH--GGRAFTTGEIIGD	330
QY	234	CNISRAKANDTLKOIVIKLRQF--ENKTIFFNHSSGGDPEIYVMSFNGCGEFPYCSTOL	292
DB	331	CNISRAKANDTLKOIVIKLRQF--ENKTIFFNHSSGGDPEIYVMSFNGCGEFPYCSTOL	390
QY	293	FNSTWNN--NTEGSSNTEG--NTITLPCRIOIIMMVEGKAMAPRISQIRIC	348
DB	391	FNSTWNN--NTEGSSNTEG--NTITLPCRIOIIMMVEGKAMAPRISQIRIC	450
QY	349	GULLTRDGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVOREK	408
DB	451	GULLTRDGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVOREK	510
QY	409	RAVGIGAVFLGFLGAGSTWGAASMTLTVOARLLLSGIYVOOQNNLLRAIEAQRM	468
DB	511	RAVGIGAVFLGFLGAGSTWGAASMTLTVOARLLLSGIYVOOQNNLLRAIEAQRM	570
QY	469	WGIRQLOARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASNSKSLDRIMNN	528
DB	571	WGIRQLOARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASNSKSLDRIMNN	630
QY	529	WEREIDNTSEIYTLIESONQOEKNEBELLELDKMASLMMWPDITNMLWY	579
DB	631	WEREIDNTSEIYTLIESONQOEKNEBELLELDKMASLMMWPDITNMLWY	681

RESULT 13  
 US-10-196-515-11  
 Sequence 11, Application US/10196515  
 Publication No. US20030091594A1  
 GENERAL INFORMATION:  
 APPLICANT: HOXIE, James A.  
 APPLICANT: LABRANCHE, Celja C.  
 APPLICANT: DOMS, Robert W.  
 APPLICANT: HOFFMAN, Trevor L.  
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
 FILE REFERENCE: Hoxie 9596-10401 (0282)  
 CURRENT APPLICATION NUMBER: US/10/196,515  
 CURRENT FILING DATE: 2002-07-16

PRIOR APPLICATION NUMBER: US/09/337,387  
 PRIOR FILING DATE: 1999-06-22  
 PRIOR APPLICATION NUMBER: US 09/317,556  
 PRIOR FILING DATE: 1999-05-24  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO 11  
 LENGTH: 856  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus type 1  
 US-10-196-515-11

Query Match 86.1%; Score 2694; DB 15; Length 856;  
 Best Local Similarity 79.4%; Pred. No. 1,7e-245;  
 Matches 517; Conservative 29; Mismatches 31; Indels 74; Gaps 6;

QY 2 EKLWVTVYGGVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVTE 61  
 DB 32 EKLWVTVYGGVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVTE 91  
 DB 62 HNNMKNMVBQWQEDIIISLDOSLKPVCVLTPLCGA----- 98  
 DB 92 NFDWMKNDWVQWQEDIIISLDOSLKPVCVLTPLCGA----- 151  
 QY 99 -----GCDTSVITQACPKISFE 115  
 DB 152 GEIKNCSEFNISTRSKVKKEAFYKLDIIPIDNDTSTLSGNTSVITQACPKISFE 211  
 QY 116 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVV 175  
 DB 212 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVV 271  
 QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKRIIRIQRGPAFTYTK-IGNMROAH 233  
 DB 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKRIIRIQRGPAFTYTK-IGNMROAH 330  
 QY 234 CNISRAKWNNTLKQIVIKLRQF-ENKTIIVNHSGGDPEIVMHSFNGGGEFFYCNSTOL 292  
 DB 331 CNISRAKWNNTLKQIVIKLRQF-ENKTIIVNHSGGDPEIVMHSFNGGGEFFYCNSTOL 390  
 QY 293 FNSWTNN--NTGSSNNTG-NTITLPCRIOIIMMGEVGMAYAPPIRQIRCSSNIT 348  
 DB 391 FNSWTNN--NTGSSNNTG-NTITLPCRIOIIMMGEVGMAYAPPIRQIRCSSNIT 450  
 QY 349 GLLTRDGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKAKRRVQREK 408  
 DB 451 GLLTRDGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKAKRRVQREK 510  
 DB 409 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQORMLQTLV 468  
 DB 511 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQORMLQTLV 570  
 QY 469 WGIQOLARVLAVERYLGDQOLLGIWCSGKLICTTAVPNNASNSKSLDRINNNMTWME 528  
 DB 571 WGIQOLARVLAVERYLGDQOLLGIWCSGKLICTTAVPNNASNSKSLDRINNNMTWME 630  
 QY 529 WEREIDNTSRLTYLIESONQOEKNEQELLEDKMASLNNWEDITNNLWY 579  
 DB 631 WEREIDNTSRLTYLIESONQOEKNEQELLEDKMASLNNWEDITNNLWY 681

RESULT 14  
 US-10-196-515-3  
 Sequence 3, Application US/10196515  
 Publication No. US20030091594A1  
 GENERAL INFORMATION:  
 APPLICANT: HOXIE, James A.  
 APPLICANT: LABRANCHE, Celja C.  
 APPLICANT: DOWS, Robert W.  
 APPLICANT: HOFFMAN, Trevor L.  
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
 FILE REFERENCE: Hoxie 9596-104U1 (0282)

CURRENT APPLICATION NUMBER: US/10/196,515  
 CURRENT FILING DATE: 2002-07-16  
 PRIOR APPLICATION NUMBER: US/09/337,387  
 PRIOR FILING DATE: 1999-06-22  
 PRIOR APPLICATION NUMBER: US 09/317,556  
 PRIOR FILING DATE: 1999-05-24  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 726  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus type 1  
 US-10-196-515-3

Query Match 85.4%; Score 2672.5; DB 15; Length 726;  
 Best Local Similarity 79.0%; Pred. No. 1,4e-243;  
 Matches 512; Conservative 31; Mismatches 32; Indels 73; Gaps 7;

QY 2 EKLWVTVYGGVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVTE 61  
 DB 32 EKLWVTVYGGVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVTE 91  
 QY 62 HNNMKNMVBQWQEDIIISLDOSLKPVCVLTPLC-----GAG----- 99  
 DB 92 NFDWMKNDWVQWQEDIIISLDOSLKPVCVLTPLC-----GAG----- 151  
 QY 100 -----GAG----- 115  
 DB 152 GEIKNCSEFNISTRSKVKKEAFYKLDIIPIDNDTSTLSGNTSVITQACPKISFE 211  
 QY 116 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVV 175  
 DB 212 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVV 271  
 QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKRIIRIQRGPAFTYTK-IGNMROAH 233  
 DB 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKRIIRIQRGPAFTYTK-IGNMROAH 330  
 QY 234 CNISRAKWNNTLKQIVIKLRQF-ENKTIIVNHSGGDPEIVMHSFNGGGEFFYCNSTOL 292  
 DB 331 CNISRAKWNNTLKQIVIKLRQF-ENKTIIVNHSGGDPEIVMHSFNGGGEFFYCNSTOL 390  
 QY 293 FNSWTNN--NTGSSNNTG-NTITLPCRIOIIMMGEVGMAYAPPIRQIRCSSNITGL 351  
 DB 391 FNSWTNN--NTGSSNNTG-NTITLPCRIOIIMMGEVGMAYAPPIRQIRCSSNITGL 448  
 QY 352 LTRPGGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKAKRRVQREK 411  
 DB 449 LTRPGGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKAKRRVQREK 508  
 QY 412 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQORMLQTLVWGI 471  
 DB 509 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQORMLQTLVWGI 568  
 QY 472 KQLOARVLAVERYLGDQOLLGIWCSGKLICTTAVPNNASNSKSLDRINNNMTWME 531  
 DB 569 KQLOARVLAVERYLGDQOLLGIWCSGKLICTTAVPNNASNSKSLDRINNNMTWME 628  
 QY 532 EIDNYSRLTYLIESONQOEKNEQELLEDKMASLNNWEDITNNLWY 579  
 DB 629 EIDNYSRLTYLIESONQOEKNEQELLEDKMASLNNWEDITNNLWY 676

RESULT 15  
 US-10-196-515-12  
 Sequence 12, Application US/10196515  
 Publication No. US20030091594A1  
 GENERAL INFORMATION:  
 APPLICANT: HOXIE, James A.  
 APPLICANT: LABRANCHE, Celja C.  
 APPLICANT: DOWS, Robert W.  
 APPLICANT: HOFFMAN, Trevor L.  
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND

/ TITLE OF INVENTION: THERAPEUTICS  
 / FILE REFERENCE: Hoxie 9596-104U1 (0282)  
 / CURRENT APPLICATION NUMBER: US/10/196,515  
 / CURRENT FILING DATE: 2002-07-16  
 / PRIOR APPLICATION NUMBER: US/09/337,387  
 / PRIOR FILING DATE: 1999-06-22  
 / PRIOR APPLICATION NUMBER: US 09/317,556  
 / PRIOR FILING DATE: 1999-05-24  
 / NUMBER OF SEQ ID NOS: 12  
 / SOFTWARE: PatentIn Ver. 2.1.  
 / SEQ ID NO: 12  
 / LENGTH: 759  
 / TYPE: PRF  
 / ORGANISM: Human immunodeficiency virus type 1  
 / US-10-196-515-12

Query Match 84.7%; Score 2651.5; DB 15; Length 759;  
 Best Local Similarity 78.4%; Pred. No. 1.5e-241;  
 Matches 508; Conservative 32; Mismatches 35; Indels 73; Gaps 6;

```

  2 EKLAATVYGVPMWKEATTTLCASDPAKAYDTEVHNWATHACVPTDPNPOEVLNVT 61
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 32 EKLAATVYGVPMWKEATTTLCASDPAKAYETEVHNWATHACVPTDPNPOEVLNVT 91
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  62 HENMKNNMVEQMOBDISLMDQSLKPCVKLTPLCGA----- 98
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 92 HENMKNNMVEQMOBDISLMDQSLKPCVKLTPLCVSLKCTDLKNDTNTSSSGRMIMK 151
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  99 -----GCDTSVITQACPKISPE 115
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  152 GEIKNCSEFNISTSKGKVKKEVAFPKLDIIPIDNDPTSYTLTSCNTSVITQACPKISPE 211
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  116 PIPHYCAPAGFAILKNDKTFNGKGPCQNVSTVOCTHGIRPVSTOLLNGSLAEVY 175
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 212 PIPHYCAPAGFAILKNNKTFNGTGPCTNVSTVOCTHGIRPVSTOLLNGSLAEVY 271
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  176 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGEIIGDIRQAH 233
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 272 IRSVNFDTAKTIIVQLNVSVEINCTKPNNTTRKIRIRIGRPGRAFVTVGK--IGNRQAH 330
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  234 CNISRAKANDTIKQIVIKLRQEP-ENKTIIVFNHSSGGDEPIVMHSGFNGGFEFFYCNSTOL 292
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 331 CNISRAKNSNTLKQIASKLREQFGNNKTIIFKQSSGGDEPIVTHSFNCGGEFFYCKSTOL 390
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  293 FNSTNNNTGSGNTEG-NTITLPCRIKQIIMMOEVEGKAMYPPIRQOIRCSNITGL 351
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 391 FNSTN--STKGSNNTGSDTITLPCRIKQIIMMOEVEKAMYPAPISQOIRCSNITGL 448
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  352 LTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIPLGVAPTCKRRVVOREKRAV 411
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 449 LTRDGINNNESEIFRPGGDMRDNRSELYKYKVKIPLGVAPTKARRVVOREKRAV 508
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  412 GIGAVFLGFLGAAGSTMGAASTLTVAQLLSGIVQOQNNLLRAIEAQOQRLQLTWGI 471
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 509 GIGAVFLGFLGAAGSTMGAASTLTVAQRLSGIVQOQNNLLRAIEAQOHLQLTWGI 568
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  472 KQLOARVLAVERYLIGDQQLLIGMGSGKLICTTAVPWNASWNSKSLDRIMNNNTMWER 531
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 569 KQLOARVLAVERYLKDQQLLIGMGSGKLICTTAVPWSASWNSKSLDRIMNNNTMWER 628
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  532 EIDNNTSEIYTLIEESONQOEKNEOBLBLDKMASIMNFDITNMLMY 579
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  DB 629 EIDNNTSLHSLIEESONQOEKNEOBLBLDKMASIMNFIISWLMY 676
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Search completed: December 12, 2003, 12:45:23  
 Job time : 26.1381 secs

**THIS PAGE BLANK (US:TO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:27:29 / Search time 13.4797 Seconds  
(without alignments)  
4130.780 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129

Sequence: 1 VEKLMVTYGVVPWKAEATTT.....ELDKWASLMMNEDITNMLMY 579

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3043	97.3	847	2 T09448	envelope glycoprot
2	3014	96.3	847	2 S13289	env protein - huma
3	2728.5	87.2	852	2 T12016	envelope glycoprot
4	2724.5	87.1	861	1 VCLJLV	env polypeptide pr
5	2719	86.9	856	1 VCLJH3	env polypeptide pr
6	2718.5	86.9	856	1 VCLJ3W	env polypeptide pr
7	2715	86.8	856	1 VCLJ3V	env polypeptide pr
8	2714	86.7	855	1 VCLJ2A	env polypeptide pr
9	2710	86.6	861	1 VCLJSC	env polypeptide pr
10	2707	86.5	854	2 S13288	env protein - huma
11	2706	86.5	843	1 H44001	env polypeptide pr
12	2701.5	86.3	851	2 S13985	env polypeptide pr
13	2696	86.2	852	1 VCLJBR	env polypeptide pr
14	2679.5	86.6	859	1 VCLJMN	env polypeptide pr
15	2628.5	84.0	861	1 VCLJKB	env polypeptide pr
16	2628.5	83.9	868	1 VCLJH4	env polypeptide pr
17	2625.5	83.9	868	1 VCLJH4	env polypeptide pr
18	2467	79.8	855	1 VCLJ2R	env polypeptide pr
19	2467	78.1	859	2 S54384	envelope polypept
20	2444.5	78.1	859	2 T01672	envelope polypept
21	2443.5	78.1	856	1 A44963	env polypeptide pr
22	2428.5	77.6	846	1 VCLJND	env polypeptide pr
23	2060	65.8	854	1 VCLJSI	env polypeptide pr
24	2014.5	64.4	443	2 C41621	env polypeptide M
25	1997.5	63.8	445	2 A41621	env polypeptide M
26	1900	60.7	506	2 A40218	envelop glycoprote
27	1858	59.4	454	2 B41621	env polypeptide D
28	1712.5	54.7	495	2 S31493	env polypeptide -
29	1688	53.9	863	2 A53034	gag polypeptide -

30	1666	53.2	877	2 S49197	envelope protein p
31	1332.5	42.6	290	2 S25940	env protein - huma
32	1331.5	41.9	297	2 S60538	envelope polypept
33	1213.5	38.8	299	2 S60554	envelope polypept
34	1208.5	38.6	299	2 S60553	envelope polypept
35	1203.5	38.5	299	2 S60552	envelope polypept
36	1199.5	38.3	299	2 S60551	envelope polypept
37	1198.5	38.3	299	2 S60528	envelope polypept
38	1198	38.3	294	2 S60545	envelope polypept
39	1196	38.2	294	2 S60525	envelope polypept
40	1195.5	38.2	299	2 S60529	envelope polypept
41	1192.5	38.1	299	2 S60521	envelope polypept
42	1190.5	38.0	299	2 S60523	envelope polypept
43	1187.5	38.0	301	2 S60548	envelope polypept
44	1187	37.9	294	2 S60524	envelope polypept
45	1182	37.8	300	2 S60546	envelope polypept

## ALIGNMENTS

## RESULT 1

T09448

envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999

C:Accession: T09448

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,

submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09448

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-847 <PAN>

A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polypeptide

Query Match	Score	Length	DB 2;	Length	847;
Best Local Similarity	89.1%	Pred. No. 2.7e-224;			
Matches	573;	Conservative	0;	Mismatches	6;
				Indels	64;
				Gaps	1;
Qy	1	VEKLMVTYGVVPWKAEATTTLCASDAKAVDEVHANWATHACPTDPNPDEVLENT	60		
Db	30	VEKLMVTYGVVPWKAEATTTLCASDAKAVDEVHANWATHACPTDPNPDEVLENT	89		
Qy	61	EHFNMMKNNMVEQOMEDIISLMDOSLKPCKVLTPLCGA-----	98		
Db	90	EHFNMMKNNMVEQOMEDIISLMDOSLKPCKVLTPLCGVLTPLCNCKDVNATNTNDSGTMERG	149		
Qy	99	-----GCDTSVITQACPKISFEP	116		
Db	150	EIKNCSFNITTSIRDEVOKEYALFYKLDVPIIDNNNTSVRLISCTSVITQACPKISFEP	209		
Qy	117	IPHYCAGAGPAILKCNKTPNGKPCCKNSVQCTHGIRPVSVSTOLLNGLAEEVVI	176		
Db	210	IPHYCAGAGPAILKCNKTPNGKPCCKNSVQCTHGIRPVSVSTOLLNGLAEEVVI	269		
Qy	177	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEGIIQDROAHQNI	236		
Db	270	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEGIIQDROAHQNI	329		
Qy	237	SRAKNDTLKOIVILRQFENKTIIVFHSSGQDEIIVHSHSGGEPFYCNSTOLFNSI	296		
Db	330	SRAKNDTLKOIVILRQFENKTIIVFHSSGQDEIIVHSHSGGEPFYCNSTOLFNSI	389		
Qy	227	WNNNTGSSNNTGNTITLPCRIKOIINMOEYGMVAPPIRGQIRGCSNITGLLTDG	356		
Db	330	WNNNTGSSNNTGNTITLPCRIKOIINMOEYGMVAPPIRGQIRGCSNITGLLTDG	449		
Qy	357	GINENGTEIFRPGGDMRDNMSLELYKTKVYKIEPLGVAPTKCKRNVVQREKAVGICAV	416		

Db 450 GINENGTETFRPGGDMEDNRSELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVGIGAV 509  
Qy 417 FLGLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 476  
Db 510 FLGLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 569  
Qy 477 RLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 536  
Db 570 RLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 629  
Qy 537 TSEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 579  
Db 630 TSEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 672

## RESULT 2

env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
Accession: S13289  
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
A:Reference number: S13288; MUID:91043044; PMID:2172833  
A:Accession: S13289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-847 <OBR>  
C:Superfamily: type E retrovirus env polyprotein

Query Match 96.3%; Score 3014; DB 2; Length 847;  
Best Local Similarity 88.5%; Pred. No. 4.4e-222;  
Matches 568; Conservative 1; Mismatches 9; Indels 64; Gaps 1;

Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61  
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90  
Qy 62 HFNMMKNMVEQOMEDIISLDQSLKPCVKLTPLC-----GAG--- 98  
Db 91 HFNMMKNMVEQOMEDIISLDQSLKPCVKLTPLC-----GAG--- 98  
Qy 99 -----GAG--- 98  
Db 151 IKNSCFNITTRIGNVQKEYALFYKLDVVPIDNNTSTRLLNSCDTSVITQACPKISFEPI 210  
Qy 118 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 177  
Db 211 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 270  
Qy 178 SDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGEIIGDIRAHCHN 237  
Db 271 SDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGEIIGDIRAHCHN 330  
Qy 238 RAKMNDTLKOIYIKLRQEPENKTIIVNHSNGGDEPIVHNSFNGGGEFFYCNSTQLNSTW 297  
Db 331 RAKMNDTLKOIYIKLRQEPENKTIIVNHSNGGDEPIVHNSFNGGGEFFYCNSTQLNSTW 390  
Qy 298 NNTGSGNNTGNTITLPCRIOIINMVEGKAMYPAPIRGOIRCSSNITGLLTRDGG 357  
Db 391 NNTGSGNNTGNTITLPCRIOIINMVEGKAMYPAPIRGOIRCSSNITGLLTRDGG 450  
Qy 358 INENGTETFRPGGDMEDNRSELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVGIGAV 417  
Db 451 INENGTETFRPGGDMEDNRSELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVGIGAV 510  
Qy 418 LGFLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 477  
Db 511 LGFLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 570  
Qy 478 VLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 537  
Db 570 VLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 630

Db 571 VLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 630  
Qy 538 SEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 579  
Db 631 SEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 672

## RESULT 3

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
Accession: T12016  
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.  
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S.  
A:Reference number: 217379; MUID:98178716; PMID:9519894  
A:Accession: T12016  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-852 <MCC>  
A:Cross-references: EMBL:U09034; NID:92351783; PIDN:AAC59271.1; PID:92351784  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 87.2%; Score 2728.5; DB 2; Length 852;  
Best Local Similarity 79.9%; Pred. No. 2.9e-200;  
Matches 517; Conservative 25; Mismatches 36; Indels 69; Gaps 4;

Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61  
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90  
Qy 62 HFNMMKNMVEQOMEDIISLDQSLKPCVKLTPLC-----GAG--- 99  
Db 91 HFNMMKNMVEQOMEDIISLDQSLKPCVKLTPLC-----GAG--- 99  
Qy 100 -----GAG--- 99  
Db 151 GEINCSFNITTRIGNVQKEYALFYKLDVVPIDNNTSTRLLNSCDTSVITQACPKISFEPI 210  
Qy 116 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 175  
Db 211 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 270  
Qy 176 IRSDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGEIIGDIRAHCHN 235  
Db 271 IRSDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGEIIGDIRAHCHN 330  
Qy 236 ISRAKNDTLKOIYIKLRQEPENKTIIVNHSNGGDEPIVHNSFNGGGEFFYCNSTQLNSTW 295  
Db 331 ISRAKNDTLKOIYIKLRQEPENKTIIVNHSNGGDEPIVHNSFNGGGEFFYCNSTQLNSTW 390  
Qy 296 TW--NNNTEGSGNNTGNTITLPCRIOIINMVEGKAMYPAPIRGOIRCSSNITGLLTRDGG 352  
Db 391 TW--NNNTEGSGNNTGNTITLPCRIOIINMVEGKAMYPAPIRGOIRCSSNITGLLTRDGG 450  
Qy 353 TRDGGINENGTETFRPGGDMEDNRSELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 412  
Db 451 TRDGGINENGTETFRPGGDMEDNRSELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 510  
Qy 413 IGALFLGLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 472  
Db 511 IGALFLGLGAGSTMGASMTLTVOARLLLSGIVOOQNLLRAIEAOQRMQLTWGIGKQQA 570  
Qy 473 QLQARVLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 532  
Db 571 QLQARVLVAERYLGDOQLLGIWCGSGKLICTTAVPMNASWSNKSIDRIINNNMTMWEREINDY 630  
Qy 533 IDNTYSEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 579  
Db 631 IDNTYSEIYTLIEESONQOEKNEOELLELDKWSASIMNMFDTNNMLY 672

RESULT 4

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N/Alternate names: coat polyprotein

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C/Accession: A03975

R/Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.

Cell 40, 9-17, 1985

A/Title: Nucleotide sequence of the AIDS virus, LAV.

A/Reference number: A90866; MUID:8509333; PMID:2981635

A/Accession: A03975

A/Molecule type: DNA

A/Residues: 1-861 <Mat>

A/Cross-reference: GB:K02013; NID:G326417; PIDN:AAE59751.1; PID:G326424

C/Genetics:

Gene: env

Performance: type B retrovirus env polyprotein

Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprotein

F.1-30/Domain: signal sequence #status predicted <SIG>

F.31-516/Product: exterior membrane glycoprotein #status predicted <EXT>

F.517-861/Product: transmembrane glycoprotein #status predicted <TM>

F.88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411,416,421,426,431,436,441,446,451,456,461,466,471,476,481,486,491,496,501,506,511,516,521,526,531,536,541,546,551,556,561,566,571,576,581,586,591,596,601,606,611,616,621,626,631,636,641,646,651,656,661,666,671,676,681,686,691,696,701,706,711,716,721,726,731,736,741,746,751,756,761,766,771,776,781,786,791,796,801,806,811,816,821,826,831,836,841,846,851,856,861,866,871,876,881,886,891,896,901,906,911,916,921,926,931,936,941,946,951,956,961,966,971,976,981,986,991,996,1001,1006,1011,1016,1021,1026,1031,1036,1041,1046,1051,1056,1061,1066,1071,1076,1081,1086,1091,1096,1101,1106,1111,1116,1121,1126,1131,1136,1141,1146,1151,1156,1161,1166,1171,1176,1181,1186,1191,1196,1201,1206,1211,1216,1221,1226,1231,1236,1241,1246,1251,1256,1261,1266,1271,1276,1281,1286,1291,1296,1301,1306,1311,1316,1321,1326,1331,1336,1341,1346,1351,1356,1361,1366,1371,1376,1381,1386,1391,1396,1401,1406,1411,1416,1421,1426,1431,1436,1441,1446,1451,1456,1461,1466,1471,1476,1481,1486,1491,1496,1501,1506,1511,1516,1521,1526,1531,1536,1541,1546,1551,1556,1561,1566,1571,1576,1581,1586,1591,1596,1601,1606,1611,1616,1621,1626,1631,1636,1641,1646,1651,1656,1661,1666,1671,1676,1681,1686,1691,1696,1701,1706,1711,1716,1721,1726,1731,1736,1741,1746,1751,1756,1761,1766,1771,1776,1781,1786,1791,1796,1801,1806,1811,1816,1821,1826,1831,1836,1841,1846,1851,1856,1861,1866,1871,1876,1881,1886,1891,1896,1901,1906,1911,1916,1921,1926,1931,1936,1941,1946,1951,1956,1961,1966,1971,1976,1981,1986,1991,1996,2001,2006,2011,2016,2021,2026,2031,2036,2041,2046,2051,2056,2061,2066,2071,2076,2081,2086,2091,2096,2101,2106,2111,2116,2121,2126,2131,2136,2141,2146,2151,2156,2161,2166,2171,2176,2181,2186,2191,2196,2201,2206,2211,2216,2221,2226,2231,2236,2241,2246,2251,2256,2261,2266,2271,2276,2281,2286,2291,2296,2301,2306,2311,2316,2321,2326,2331,2336,2341,2346,2351,2356,2361,2366,2371,2376,2381,2386,2391,2396,2401,2406,2411,2416,2421,2426,2431,2436,2441,2446,2451,2456,2461,2466,2471,2476,2481,2486,2491,2496,2501,2506,2511,2516,2521,2526,2531,2536,2541,2546,2551,2556,2561,2566,2571,2576,2581,2586,2591,2596,2601,2606,2611,2616,2621,2626,2631,2636,2641,2646,2651,2656,2661,2666,2671,2676,2681,2686,2691,2696,2701,2706,2711,2716,2721,2726,2731,2736,2741,2746,2751,2756,2761,2766,2771,2776,2781,2786,2791,2796,2801,2806,2811,2816,2821,2826,2831,2836,2841,2846,2851,2856,2861,2866,2871,2876,2881,2886,2891,2896,2901,2906,2911,2916,2921,2926,2931,2936,2941,2946,2951,2956,2961,2966,2971,2976,2981,2986,2991,2996,3001,3006,3011,3016,3021,3026,3031,3036,3041,3046,3051,3056,3061,3066,3071,3076,3081,3086,3091,3096,3101,3106,3111,3116,3121,3126,3131,3136,3141,3146,3151,3156,3161,3166,3171,3176,3181,3186,3191,3196,3201,3206,3211,3216,3221,3226,3231,3236,3241,3246,3251,3256,3261,3266,3271,3276,3281,3286,3291,3296,3301,3306,3311,3316,3321,3326,3331,3336,3341,3346,3351,3356,3361,3366,3371,3376,3381,3386,3391,3396,3401,3406,3411,3416,3421,3426,3431,3436,3441,3446,3451,3456,3461,3466,3471,3476,3481,3486,3491,3496,3501,3506,3511,3516,3521,3526,3531,3536,3541,3546,3551,3556,3561,3566,3571,3576,3581,3586,3591,3596,3601,3606,3611,3616,3621,3626,3631,3636,3641,3646,3651,3656,3661,3666,3671,3676,3681,3686,3691,3696,3701,3706,3711,3716,3721,3726,3731,3736,3741,3746,3751,3756,3761,3766,3771,3776,3781,3786,3791,3796,3801,3806,3811,3816,3821,3826,3831,3836,3841,3846,3851,3856,3861,3866,3871,3876,3881,3886,3891,3896,3901,3906,3911,3916,3921,3926,3931,3936,3941,3946,3951,3956,3961,3966,3971,3976,3981,3986,3991,3996,4001,4006,4011,4016,4021,4026,4031,4036,4041,4046,4051,4056,40

[illegible]

Dh 571 WGIKOLARILAVERYLKDQQLLIGWCSGKLICTTAVPMNASWSNKSLEIMNMNTWME 630  
Qy 529 MEREIDNTSEIYTLIESONQOEKNBOELLDKMASLWMPDITWMLWY 579  
Dh 631 WDRINNTSLIHLIESONQOEKNBOELLDKMASLWMPDITWMLWY 681

## RESULT 6

VCLJ3M  
env polypotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)  
N:Contamin: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 07-Nov-1997  
C:Accession: A24774  
R:Starch: B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.  
Cell 45; 637-648; 1986  
A:Title: Identification and characterization of conserved and variable regions in the env  
A:Reference number: A24774; MUID:86218077; PMID:2423250  
A:Accession: A24774  
A:Molecule type: DNA  
A:Residues: 1-856 <STR>  
A:Cross-references: GB:K03455; GB:M38432; NID:91906382  
C:Genetics:

A:Gene: env  
C:Superfamily: type B retrovirus env polypotein  
C:Keywords: coat protein; glycoprotein; polypotein; transmembrane protein  
F:1-29/Domin: signal sequence #status predicted <SIG>  
F:30-501/Product: coat protein gp120 #status predicted <GP1>  
F:502-847/Product: coat protein gp41 #status predicted <GP2>  
F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,455  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Aam) (covalent) #status predict

Query Match 86.9%; Score 2718.5; DB 1; Length 856;  
Best Local Similarity 79.9%; Pred. No. 1-76-199;  
Matches 522; Conservative 24; Mismatches 32; Indels 75; Gaps 7;

Qy 1 VEKLVTVYGVVPWKEATTTLFCASDAKAYDTEVHVMATHACVPTDPNQEVLNVT 60  
Dh 30 VEQLMTVYGVVPWKEATTTLFCASDAKAYDTEVHVMATHACVPTDPNQEVLNVT 89  
Qy 61 EHPMMKNNWEOQNEEDIIISLMDQSLKPCVLTPLC-----GAG-- 99  
Dh 90 EHPMMKNNWEOQNEEDIIISLMDQSLKPCVLTPLC-----GAG-- 99  
Qy 100 -----CDTSVITQACPKIS 113  
Dh 150 KNCSEFNITSIKDKVKEAFYKLDVPIKSNDSITRYRLHCHNTSVITQACPKIS 209  
Qy 114 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGLAEE 173  
Dh 210 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGLAEE 269  
Qy 174 VVIRSDFTNNAKTIIVOLKESVEINCRPNNT-RKSIHIGRPAPFTTGIIIGDIROA 232  
Dh 270 IVIRSENFDAKTIIVOLKESVEINCRPNNT-RKSIHIGRPAPFTTGIIIGDIROA 328  
Qy 233 HCNISRAKNDTTLKOIYIKLREOFENKTIIVNHSNGGDEIIVHSPNGGGEFFYCNSTOL 292  
Dh 329 HCNISRAKNDTTLKOIYIKLREOFENKTIIVNHSNGGDEIIVHSPNGGGEFFYCNSTOL 388  
Qy 293 FNSFTN--NNTGSSNTE--GNTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 347  
Dh 389 FNSFTNVTGISSTEGNNTTEENGDTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 448  
Qy 348 TGLLLTDDGNGENGTIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 407  
Dh 449 TGLLLTDDGNGSSRREIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 508  
Qy 408 KRAVGI-IGAVFLGFLGAAGSTGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQMLQ 466  
Dh 509 KRAVGI-IGAVFLGFLGAAGSTGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQMLQ 568  
Qy 467 TWVGIKOLQARVLAVERYLDQQLLIGWCSGKLICTTAVPMNASWSNKSLEIMNMNTW 526

Dh 569 TWVGIKOLQARVLAVERYLDQQLLIGWCSGKLICTTAVPMNASWSNKSMDQIMNMNTW 628  
Qy 527 MEREIDNTSEIYTLIESONQOEKNBOELLDKMASLWMPDITWMLWY 579  
Dh 629 MEREIDNTSEIYTLIESONQOEKNBOELLDKMASLWMPDITWMLWY 681

## RESULT 7

VCLJLV  
env polypotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N:Alternate names: coat polypotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C>Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03974  
R:Mesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313; 450-458; 1985  
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi  
A:Reference number: A93355; MUID:85111157; PMID:2982104  
A:Accession: A03974  
A:Molecule type: DNA  
A:Residues: 1-856 <MSE>  
A:Cross-references: GB:K02083; NID:9555008; PIDN:AAB59873.1; PID:9328559  
C:Genetics:

A:Gene: env  
C:Superfamily: type B retrovirus env polypotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypote  
F:1-30/Domin: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <TM>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Aam) (covalent) #status predict

Query Match 86.8%; Score 2715; DB 1; Length 856;  
Best Local Similarity 79.9%; Pred. No. 3-26-199;  
Matches 520; Conservative 27; Mismatches 30; Indels 74; Gaps 6;

Qy 2 EKLMTVYGVVPWKEATTTLFCASDAKAYDTEVHVMATHACVPTDPNQEVLNVT 61  
Dh 32 EKLMTVYGVVPWKEATTTLFCASDAKAYDTEVHVMATHACVPTDPNQEVLNVT 91  
Qy 62 HPMKNNWEOQNEEDIIISLMDQSLKPCVLTPLC-----GAG-- 98  
Dh 92 HPMKNNWEOQNEEDIIISLMDQSLKPCVLTPLC-----GAG-- 98  
Qy 99 -----CDTSVITQACPKIS 115  
Dh 152 GEIKNCSFNISTIRKQVKEAFYKLDIIPIDNDSITRYRLHCHNTSVITQACPKIS 211  
Qy 116 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGLAEE 175  
Dh 212 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGLAEE 269  
Qy 176 IRSDFNTNNAKTIIVOLKESVEINCRPNNT-RKSIHIGRPAPFTTGIIIGDIROA 233  
Dh 272 IRSDFNTNNAKTIIVOLKESVEINCRPNNT-RKSIHIGRPAPFTTGIIIGDIROA 328  
Qy 234 CNISRAKNDTTLKOIYIKLREOF-ENKTIIVNHSNGGDEIIVHSPNGGGEFFYCNSTOL 292  
Dh 331 CNISRAKNDTTLKOIYIKLREOF-ENKTIIVNHSNGGDEIIVHSPNGGGEFFYCNSTOL 388  
Qy 293 FNSFTN--NNTGSSNTE--GNTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 347  
Dh 391 FNSFTNVTGISSTEGNNTTEENGDTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 448  
Qy 349 GLLLTDDGNGENGTIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 408  
Dh 451 GLLLTDDGNGNNSEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 510  
Qy 409 RAVGIGAVFLGFLGAAGSTGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQMLQ 468  
Dh 511 RAVGIGAVFLGFLGAAGSTGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQMLQ 570

QY 469 WGIKQIQAARVLAVERLYDQOOLLGIMGSGSKLICCTAIVPWNASWSNKSIDRIINNNNTWNE 528  
 Db 571 WGIKQIQAARVLAVERLYDQOOLLGIMGSGSKLICCTAIVPWNASWSNKSIDRIINNNNTWNE 630  
 QY 529 WEREIDNTYSEIYTLIEESQNOOEKNEOBELLEIDKASLMMNPDITNMLMY 579  
 Db 631 WEREIDNTYSLIHSIEESQNOOEKNEOBELLEIDKASLMMNPDITNMLMY 681

## RESULT 8

VCLJ32

env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A&gt;Note: host Homo sapiens (man)

C&gt;Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A03976

R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Sh

Science 227, 484-492, 1985

A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).

A:Reference number: A04003; MUID:85090453; PMID:2578227

C:Accession: A03976

A:Molecule type: DNA

A:Residues: 1-855 &lt;SAN&gt;

A:Cross-references: GB:K02007; MID:G328658; PIDN:AAB59882.1; PID:G328666

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

F:1-30/Domain: signal sequence #status predicted &lt;SIG&gt;

F:31-509/Product: transmembrane glycoprotein #status predicted &lt;TM&gt;

F:510-855/Product: transmembrane glycoprotein #status predicted &lt;TM&gt;

F:87,129,140,158,184,190,200,244,265,292,298,304,334,341,358,364,388,394,400,408,445,458

F:610,624,636,615/binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 86.7%; Score 2714; DB 1; Length 855;

Best Local Similarity 79.3%; Pred. No. 3,8e-199;

Matches 518; Conservative 27; Mismatches 30; Indels 78; Gaps 7;

QY 2 EKLWTVVYGVVPMKEATTTLCASDARAYDTEVHNWATHACVPTDNPQEVLENTE 61  
 Db 31 EKLWTVVYGVVPMKEATTTLCASDARAYDTEVHNWATHACVPTDNPQEVLENTE 90  
 QY 62 HFNMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98  
 Db 91 HFNMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150  
 QY 99 -----GCDTSVITQACPK 111  
 Db 151 EIKNSFNITTSIRDKIQENALFRLDVVPIDNASTTNTNRYRLHICNRSVITQACPK 210  
 QY 112 ISFEPIPIHYCAPAGPAILKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAE 171  
 Db 211 VSFEPIPIHYCTPAGPAILKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAE 270  
 QY 172 BEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 231  
 Db 271 BEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 330  
 QY 232 AHCNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEIIVHVSFNGCGEFPYCNS 290  
 Db 331 AHCNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEIIVHVSFNGCGEFPYCNS 390  
 QY 391 QLFNFTNNTNTEGSSNNTBGN--TITLPCRIKQIINMOQEVGKAMYPAPPIGQIRGSSNT 348  
 Db 447 QLFNFTNNTNTEGSSNNTBGN--TITLPCRIKQIINMOQEVGKAMYPAPPIGQIRGSSNT 447  
 QY 448 GLLLTTRDGIN--ENGTEIFRPGGDMRDMRSELKYKVKVIEPLGVAATTKKRRVORE 407  
 Db 507 GLLLTTRDGIN--ENGTEIFRPGGDMRDMRSELKYKVKVIEPLGVAATTKKRRVORE 507  
 QY 468 KRAVGI--GAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYVOQNNLLRAIEAOQL 466  
 Db 466 KRAVGI--GAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYVOQNNLLRAIEAOQL 466

Db 508 KRAVGI--GAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYVOQNNLLRAIEAOQL 567  
 QY 467 TWGIKQIQAARVLAVERLYDQOOLLGIMGSGSKLICCTAIVPWNASWSNKSIDRIINNNNTW 526  
 Db 568 TWGIKQIQAARVLAVERLYDQOOLLGIMGSGSKLICCTAIVPWNASWSNKSIDRIINNNNTW 627  
 QY 527 MEMEREIDNTYSEIYTLIEESQNOOEKNEOBELLEIDKASLMMNPDITNMLMY 579  
 Db 628 MEMEREIDNTYSLIHSIEESQNOOEKNEOBELLEIDKASLMMNPDITNMLMY 680

## RESULT 9

VCLJ3C

env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A&gt;Note: host Homo sapiens (man)

C&gt;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997

C:Accession: B28922

R:Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta

Virology 164, 531-536, 1988

A:Title: Envelope sequences of two new United States HIV-1 isolates.

A:Reference number: A28922; MUID:86219542; PMID:3365091

A:Molecule type: DNA

A:Residues: 1-661 &lt;GUR&gt;

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote

F:1-29/Domain: signal sequence #status predicted &lt;SIG&gt;

F:30-861/Product: env polyprotein #status predicted &lt;EP&gt;

F:87,129,135,140,143,159,163,187,198,224,241,262,276,285,301,302,333,340,356,362,386,396

Query Match 86.6%; Score 2710; DB 1; Length 861;

Best Local Similarity 79.2%; Pred. No. 7.7e-199;

Matches 521; Conservative 25; Mismatches 30; Indels 82; Gaps 8;

QY 2 EKLWTVVYGVVPMKEATTTLCASDARAYDTEVHNWATHACVPTDNPQEVLENTE 61  
 Db 31 EKLWTVVYGVVPMKEATTTLCASDARAYDTEVHNWATHACVPTDNPQEVLENTE 90  
 QY 62 HFNMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98  
 Db 91 HFNMMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150  
 QY 99 -----GCDTSVITQACPK 113  
 Db 151 KMEGEMTNCSEFNITTSIRSKQKEYALFYKLDVVPIDNDTSTYLINCNSTSVITQACPK 210  
 QY 114 FEPIPIHYCAPAGPAILKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAE 173  
 Db 211 FEPIPIHYCA--RWFAILCNKKNKFTNGTCTNVSTVOCTHGIRPVVSTOLLNGSLAE 269  
 QY 174 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 231  
 Db 270 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 329  
 QY 232 AHCNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEIIVHVSFNGCGEFPYCNS 291  
 Db 330 AHCNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEIIVHVSFNGCGEFPYCNS 389  
 QY 390 LFNSTNNNTNTEGSSNNTBGN--TITLPCRIKQIINMOQEVGKAMYPAPPIGQIRGSSNT 347  
 Db 448 LFNSTNNNTNTEGSSNNTBGN--TITLPCRIKQIINMOQEVGKAMYPAPPIGQIRGSSNT 448  
 QY 449 TGLLTTRDGIN--ENGTEIFRPGGDMRDMRSELKYKVKVIEPLGVAATTKKRRVORE 402  
 Db 508 TGLLTTRDGIN--ENGTEIFRPGGDMRDMRSELKYKVKVIEPLGVAATTKKRRVORE 508  
 QY 403 VVIREKRAVG--IGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYVOQNNLLRAIEAO 461  
 Db 568 VVIREKRAVG--IGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYVOQNNLLRAIEAO 568

```
Qy 462 RMLQLTWGIKQIOLARVLAVERYLGDQQLIGIWCSCGLICTTAVPWNASWNSKSLDRIM 521
Db 569 HLLQLTWGIKQIOLARVLAVERYLGDQQLIGIWCSCGLICTTAVPWNASWNSKSLDRIM 628
Qy 522 NNMWMEWEIREIDNYTSEIYTLIEESQNOEKNEQELLEDKMASLWNPFDITWLMWY 579
Db 629 GNMWMEWEIREIDNYTSEIYTLIEESQNOEKNEQELLEDKMASLWNPFDITWLMWY 686

RESULT 10
S13288
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13288
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A:Title: HIV-1 trioplem for mononuclear phagocytes can be determined by regions of gp120
Reference number: S13288; MUID:91043044; PMID:2172833
Session: S13288
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-854 <OBR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 86.5%; Score 2707; DB 2; Length 854;
Best Local Similarity 79.8%; Pred. No. 1.3e-198;
Matches 518; Conservative 25; Mismatches 34; Indels 72; Gaps 6;

Qy 2 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 61
Db 32 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 91
Qy 62 HFNWKNMWEQNOEDIIISLMDOSLKPCVKLTPLCGA----- 98
Db 92 HFNWKNMWEQNOEDIIISLMDOSLKPCVKLTPLCGA----- 98
Qy 99 -----GCDTSVITQCPKISFEP 117
Db 152 GEIKNCSFNISTSRDKVQKEAFYKLDIVPINDTSYRLISCNISVITQCPKVSFEP 211
Qy 118 PIHYCAPAGFAILKCNKTFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEVIR 177
Db 212 PIHYCAPAGFAILKCNKTFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEVIR 271
Qy 178 SDNTNNAKTIIVOLKESVEINCTRPNNNTKSIHI--GGRAPYTTGELIGDIRAHCHN 235
Db 272 SANFTDAKTIIVOLNTSVEINCTRPNNNTKSIIRIORGPRAFTVIGK--IGNRROAHCHN 330
Qy 236 ISRAKMDTLKOIYIKLREOF--ENKTIYFNHSSGGDEPIVHSHFNCGEFFYCNSTOLF 294
Db 331 ISRAKMDTLKOIYIKLREOF--ENKTIYFNHSSGGDEPIVHSHFNCGEFFYCNSTOLF 390
Qy 295 STWNN--NTEGSNNTEG--NTITLPCRKOIINNMQEVGKAMVAPPIRGQIRCSNITGL 350
Db 391 STWNN--NTEGSNNTEG--NTITLPCRKOIINNMQEVGKAMVAPPIRGQIRCSNITGL 450
Qy 351 LITRDGGINENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 410
Db 451 LITRDGGINENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 510
Qy 411 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 470
Db 511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 570
Qy 471 IKOLQAEVLAVERYLGDQQLIGIWCSCGLICTTAVPWNASWNSKSLDRIMNNTWMEWE 530
Db 571 IKOLQAEVLAVERYLGDQQLIGIWCSCGLICTTAVPWNASWNSKSLDRIMNNTWMEWE 630
Qy 531 REIDNYTSEIYTLIEESQNOEKNEQELLEDKMASLWNPFDITWLMWY 579
Db 631 REIDNYTSEIYTLIEESQNOEKNEQELLEDKMASLWNPFDITWLMWY 679
```

```
RESULT 11
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
C:Accession: H44001
R:L.I. Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
J. Virol. 66, 6587-6600, 1992
A:Title: Complete nucleotide sequence, genome organization, and biological properties of
A:Reference number: A44001; MUID:93021387; PMID:1404605
A:Accession: H44001
A:Molecule type: DNA
A:Residues: 1-843 <LIY>
A:Cross-references: GB:M93258
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:19-35/Region: hydrophobic
F:30-489/Product: coat protein gp120 #status predicted <GP1>
F:490-843/Product: coat protein gp41 #status predicted <GP2>
F:499-515/Region: hydrophobic
F:673-689/Region: hydrophobic
F:728-755/Domain: transmembrane #status predicted <TMN>
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435,
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435,

Query Match 86.5%; Score 2706; DB 1; Length 843;
Best Local Similarity 79.8%; Pred. No. 1.5e-198;
Matches 513; Conservative 25; Mismatches 36; Indels 68; Gaps 5;

Qy 2 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 61
Db 31 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 90
Qy 62 HFNWKNMWEQNOEDIIISLMDOSLKPCVKLTPLCGA----- 98
Db 91 HFNWKNMWEQNOEDIIISLMDOSLKPCVKLTPLCGA----- 98
Qy 99 -----GCDTSVITQCPKISFEP 118
Db 151 EIKNCSFNISTSRDKVQKEAFYKLDIVPINDTSYRLISCNISVITQCPKVSFEP 210
Qy 119 IHYCAPAGFAILKCNKTFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEVIR 178
Db 211 IHYCAPAGFAILKCNKTFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEVIR 270
Qy 179 DNFTNNAKTIIVOLKESVEINCTRPNNNTKSIHIGGRAPYTTGELIGDIRAHCHN 238
Db 271 DNFTNNAKTIIVOLKESVEINCTRPNNNTKSIHIGGRAPYTTGELIGDIRAHCHN 330
Qy 239 AKMNDTLKOIYIKLREOF--ENKTIYFNHSSGGDEPIVHSHFNCGEFFYCNSTOLF--TW 388
Db 331 AKMNDTLKOIYIKLREOF--ENKTIYFNHSSGGDEPIVHSHFNCGEFFYCNSTOLF--TW 388
Qy 351 LITRDGGINENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 417
Db 411 LITRDGGINENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 510
Qy 417 INENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 477
Db 477 INENGTEIFRPGGDMRDNRSELKYKIVYKIEPLGVAFTCKRRVYQREKRA 570
Qy 477 LGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 530
Db 530 LGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 630
Qy 530 LGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 579
Db 579 LGFLGAAGSTWGAASMTLTVQARLLLSGIYQOQNNLLRAIEAOQRLQLTVWG 679
```





N.Containers: coat protein gp120; coat protein gp32  
C.Species: human immunodeficiency virus type 1, HIV-1  
A.Note: host Homo sapiens (man)  
C.Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 01-Mar-1996  
C.Accession: B42995  
R.Shimizu, H.; Hasabe, F.; Teuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.  
Virology 189, 534-546, 1992  
A.Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated  
A.Reference number: A42995; M0ID:92351552; PMID:1122587  
A.Accession: B42995  
A.Molecule type: mRNA  
A.Residues: 1-729 <SHI>  
A.Cross-references: GB:S41266; GB:D01206  
C.Genetics:  
A.Gene: env  
C.Superfamily: type E retrovirus env polypeptide  
C.Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F.I-689/Domain: extracellular #status predicted <EXT>  
F.I-1-33/Domain: signal sequence #status predicted <SIG>  
F.I-1-33/Region: hydrophobic #status predicted  
F.I-34-517/Product: coat protein gp120 #status predicted <CP1>  
F.I-34-517/Region: cleavage processing #status predicted  
F.I-518-729/Product: coat protein gp32 #status predicted <CP2>  
F.I-518-533/Region: hydrophobic #status predicted  
F.I-690-711/Domain: transmembrane #status predicted <TM>  
F.I-712-729/Domain: intracellular #status predicted <INT>  
F.I-93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414

Query Match 84.0%; Score 2628.5; DB 1; Length 729;  
Best Local Similarity 75.9%; Pred. No. 1e-192;  
Matches 494; Conservative 33; Mismatches 51; Indels 73; Gaps 2;

QY 2 EKLTWTVVVYGVPMKEATTTLFCASDAKAYDTEVNHNWATHACVPDPNQEVLNVTE 61  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 37 EQLTWTVVVYGVPMKEATTTLFCASDAKAYDTFAHWNWATHACVPDPNQEVLNVTE 96  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 62 HNNMKNNWVEOMOEIIISLMDOSLRPCVKLPGLCGA----- 98  
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 97 NFNMMKNWVEOMHENIISLMDOSLRPCVKLPGLCVTLHCTDLNTNNNSIEKKGE 156  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 99 -----GCDTSVTQACPIS 113  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 157 IKNGCFNVTNRDKYOKRYALFYKLDVYPINDNNSTNYCLISDTSVITDACRVYS 216  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 114 FEPIPIHYCAPAFAILKCNKDIFNGKGPCKNVSVQCTHGIRPVYSTOLLNLSLAEE 173  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 217 FEPIPIHYCTPAGFALKCNKKTFFNGTGPCKNVSVQCTHGIRPVYSTOLLNLSLAEG 276  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 174 VWRSDNFNNNAKTIIVQLKESYEINCTRPNNTRKSINHGRPAFTTGTEIIDIRQA 233  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 277 VWRSENFNDNKTIIVQLNEITYKINCIRPNNTKRKYVMPSKVYTTTEIIDIRQA 336  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 234 CNISAKNMDDLKOVIYKLREOFENKTIYFNHSSGGDPETVMHSFNGGFEFFYCSNQLF 293  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 337 CNISAEWMKTFEQLANKRKPFENKTIYFNQSSGGDPETVMHFNFGGFEFFYCDSSOLF 396  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 294 NSTANNNTGSGNNTEGNITTLPCRIRKIINMEOVGKAMAPPIRGDIRCSSNTIGLLT 353  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 397 NSTHLSTNGVMWNGTGENTTLPCRIRKIYNMDEVGKAWAPPIRGDIRCSSNTIGLLT 456  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 354 RDGINENG-----EIFPPGGDMEDMRSELKYKRVKIEPLGVAPTYCKRRVOREK 408  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 457 RDGGTNONTNNTSLEIFRPGGGDMDRMSRELKYKRVKIEPLGVAPTAKRRVOREK 516  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 409 RAVVGGAVALGFLGAAGSTMGAASMTLVQARLLSGIVOOQNLLRAIEAOQMLDLTV 468  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 517 RAVVGGAVALGFLGAAGSTMGAASMTLVQARLLSGIVOOQNLLRAIDQQHLDLTV 576  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 469 WGIKLOARVLAVERYLGDQQLLGICGSGKLICTAVPWNASMSNKLDRINNMTWME 528  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 577 WGIKLOARVLAVERYLKDQQLGMIGWGSGKFICTTAVPNTSMNSKSPFEIMDNMTWME 636  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 529 WEREIDNTYTSLEYTLLESQNOQEKNEDELILDKNASLMNWPDITNWLMY 579



Db 637 WERINNYTNLYNLIBESQOQEKNEODLALDKWDSLNNWFSITKMLTY 687

Search completed: December 12, 2003, 12:33:33  
Job time : 15.4797 secs

**THIS PAGE BLANK (US-770)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:26:14 ; Search time 29.7807 Seconds  
(without alignments)  
5017.066 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129  
Sequence: 1 VKELWVTYYGVPMKEATT.....ELDKNASLWNPEDITWLMY 579

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seque, 258052604 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP rvirus:\*  
16: SP bacteriap:\*  
17: SP archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3043	97.3	847	15	Q75760 human immun
2	2824.5	90.3	838	15	Q03806 human immun
3	2814.5	89.9	854	15	Q78225 human immun
4	2800.5	89.5	855	15	Q03805 human immun
5	2789.5	89.1	848	15	Q74999 human immun
6	2784.5	89.0	854	15	Q40222 human immun
7	2780.5	88.9	852	15	Q92761 human immun
8	2777	88.8	853	15	Q03811 human immun
9	2770	88.5	854	15	Q92762 human immun
10	2769	88.3	863	15	Q56110 human immun
11	2759	88.2	851	15	Q77989 human immun
12	2752.5	88.0	846	15	O56562 human immun
13	2752	87.8	847	15	Q9PWX7 human immun
14	2747	87.8	843	15	Q8AR21 human immun
15	2745	87.7	861	15	Q70150 human immun
16	2745	87.7	861	15	Q9WJVS human immun

17	2743.5	87.7	853	15	O56563 human immun
18	2742	87.6	853	15	O56108 human immun
19	2741.5	87.6	854	15	O92875 human immun
20	2740.5	87.6	843	15	Q70008 human immun
21	2738	87.5	859	15	Q72940 human immun
22	2738	87.5	860	15	Q9E157 human immun
23	2736	87.4	849	15	Q77368 human immun
24	2735.5	87.4	860	15	O9YP50 human immun
25	2734.5	87.4	868	15	O9MJWS human immun
26	2734	87.4	861	15	O9E527 human immun
27	2733	87.3	861	15	O9E523 human immun
28	2732.5	87.3	841	15	Q9QKJ1 human immun
29	2732	87.3	838	15	O9Q6V0 human immun
30	2730.5	87.3	735	15	O9IWI6 human immun
31	2730.5	87.3	852	15	O56567 human immun
32	2729.5	87.2	856	15	Q72993 human immun
33	2728.5	87.2	852	15	O41883 human immun
34	2726.5	87.1	854	15	O56112 human immun
35	2726	87.1	863	15	O42031 human immun
36	2725	87.1	851	15	Q73304 human immun
37	2724	87.1	857	15	Q8UL70 human immun
38	2723.5	87.0	860	15	Q9DH01 human immun
39	2723	87.0	855	15	Q8Q367 human immun
40	2723	87.0	863	15	Q9WJU8 human immun
41	2722.5	87.0	850	15	Q70003 human immun
42	2722	87.0	752	15	Q70604 human immun
43	2722	87.0	752	15	O70605 human immun
44	2721	87.0	853	15	Q9J023 human immun
45	2720	86.9	851	15	Q8Q852 human immun

## ALIGNMENTS

### RESULT 1

ID Q75760 PRELIMINARY; PRT; 847 AA.  
AC Q75760;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRF;  
RX MEDLINE=91043044; PubMed=3646751;  
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K.,  
RA Koyanagi Y., Miles S., Mitsuya R.T., Merrill J.E., Vinters H.V.,  
RA Chen I.S.;  
RT "Dual infection of the central nervous system by AIDS viruses with  
RT distinct cellular tropisms.";  
RL Science 236:819-822(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRF;  
RX MEDLINE=91043044; PubMed=2172833;  
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K.,  
RA Zack J.A., Chen I.S.;  
RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions  
RT of gp120 outside the CD4-binding domain.";  
RL Nature 348:69-73(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JRF;  
RX MEDLINE=92092169; PubMed=1684385;  
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;  
RT "HIV-1 env sequence variation in brain tissue of patients with AIDS-  
RT related neurologic disease.";  
RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).  
RN [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN=JREL;  
 RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.,  
 RA Kojanagi Y., Namazie A., Zhao J., Diagne A., Idler K.;  
 RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63632; AAB05604.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SO SEQUENCE 847 AA; 96160 MW; 022D5F24E04FB29F CRC64;

Query Match 97.3%; Score 3043; DB 15; Length 847;  
 Best Local Similarity 89.1%; Pred. No. 5e-252;  
 Matches 573; Conservative 0; Mismatches 6; Indels 64; Gaps 1;

QY 1 VEKLMVTYYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVVLENVT 60  
 30 VEKLMVTYYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVVLENVT 89  
 QY 61 EHFNMKNMWEQWEDIIISLMDSLKPCVKLTPLCGA----- 98  
 90 EHFNMKNMWEQWEDIIISLMDSLKPCVKLTPLCGA----- 149  
 QY 99 -----GCDTSVITQACPKISFEP 116  
 150 EIKNCSFNTTISINDEQKEYALFYKLDVPIDNNNTSYRLISCDTSVITQACPKISFEP 209  
 Db 150 EIKNCSFNTTISINDEQKEYALFYKLDVPIDNNNTSYRLISCDTSVITQACPKISFEP 209  
 QY 117 IPIHYCAPAGPAILKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI 176  
 210 IPIHYCAPAGPAILKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI 269  
 Db 210 IPIHYCAPAGPAILKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI 269  
 QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI 236  
 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI 329  
 Db 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI 329  
 QY 237 SRAKNDTLKOIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 296  
 330 SRAKNDTLKOIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 389  
 Db 330 SRAKNDTLKOIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 389  
 QY 297 WNNTEGSGNTGNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 356  
 390 WNNTEGSGNTGNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 449  
 Db 390 WNNTEGSGNTGNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 449  
 QY 357 GINENGTEIFRPGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAV 416  
 450 GINENGTEIFRPGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAV 509  
 Db 450 GINENGTEIFRPGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAV 509  
 QY 417 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQA 476  
 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQA 569  
 Db 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQA 569  
 QY 477 RVLAVERLYGQOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 536  
 570 RVLAVERLYGQOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 629  
 Db 570 RVLAVERLYGQOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 629  
 QY 537 TSEIYTLIEESONOEKNEOELELDKMASLMMWFDTNNMLMY 579  
 630 TSEIYTLIEESONOEKNEOELELDKMASLMMWFDTNNMLMY 672  
 Db 630 TSEIYTLIEESONOEKNEOELELDKMASLMMWFDTNNMLMY 672

RESULT 2  
 003806 PRELIMINARY; PRT; 838 AA.  
 AC 003806;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Env protein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Comole E.,  
 RA Beaver B.;  
 RA Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M68894; AAA44196.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 SO SEQUENCE 838 AA; 95329 MW; 11044B1CA10CFB4D CRC64;

Query Match 90.3%; Score 2824.5; DB 15; Length 838;  
 Best Local Similarity 83.7%; Pred. No. 2.8e-233;  
 Matches 530; Conservative 20; Mismatches 28; Indels 55; Gaps 2;

QY 2 EKLMTVYYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVVLENVT 61  
 31 EKLMTVYYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVVLENVT 90  
 Db 31 EKLMTVYYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVVLENVT 90  
 QY 62 EHFNMKNMWEQWEDIIISLMDSLKPCVKLTPLCGA-----GAG----- 99  
 91 EHFNMKNMWEQWEDIIISLMDSLKPCVKLTPLCGA-----GAG----- 150  
 QY 100 -----GCDTSVITQACPKISFEP 126  
 151 TNIRGKQKEYALFYELDIPIPNKNDYRLISGNTSVITQACPKISFEP 210  
 Db 151 TNIRGKQKEYALFYELDIPIPNKNDYRLISGNTSVITQACPKISFEP 210  
 QY 127 FALIKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI RSDNFTNNAK 186  
 211 FALIKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI RSDNFTNNAK 270  
 Db 211 FALIKCNKNTNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAEEVVI RSDNFTNNAK 270  
 QY 187 TIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI SRAKNDTLK 246  
 271 TIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI SRAKNDTLK 330  
 Db 271 TIIVOLKESVEINCTRPNNNTRKSHIGPRAFTTGEIIGDIRQAHNI SRAKNDTLK 330  
 QY 247 QIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 306  
 331 KIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 390  
 Db 331 KIVIKLREQENKTIIVFNHSSGGDEPEIWMHSPNCGEFPYCNSTQLFNST 390  
 QY 307 TEGNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 366  
 391 VENNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 450  
 Db 391 VENNTITLPCRIKOIIMMOEYKAMVAPPIRGQIRCSSNITGLLTRDG 450  
 QY 367 RFGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAVFLGFLGAAGS 426  
 451 RFGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAVFLGFLGAAGS 510  
 Db 451 RFGGDMRDMRSELYKVKYKIEPLGVAFTKARRVQREKRAVGIGAVFLGFLGAAGS 510  
 QY 427 TGAASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQARVLAVERLYG 486  
 511 TGAASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQARVLAVERLYG 570  
 Db 511 TGAASMTLTVQARLLLSGIYQOQNNLRAIEAQORMQLTVMGIKOQARVLAVERLYG 570  
 QY 487 DOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 546  
 571 DOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 630  
 Db 571 DOQLLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTWMEWEIEDNY 630  
 QY 547 SONOEKNEOELELDKMASLMMWFDTNNMLMY 579  
 631 SONOEKNEOELELDKMASLMMWFDTNNMLMY 663  
 Db 631 SONOEKNEOELELDKMASLMMWFDTNNMLMY 663

RESULT 3  
 078225 PRELIMINARY; PRT; 854 AA.  
 AC 078225;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Env protein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovid viruses; Retroviridae; Lentivirus.



```

DE Envelope glycoprotein gp160.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JR-CSF;
RX MEDLINE=97063128; PubMed=8906996;
RA Klasse P.J., Boyd M.T., Weiss R.A., Schulz T.F.;
RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing
RT variant of HIV type 1 JR-CSF that infects a range of T cell lines.";
RL AIDS Res. Hum. Retroviruses 12:347-350(1996).
DR EMBL; U45960; AAB60591.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KV AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT VARIANT 164 164 N -> D.
FT VARIANT 183 183 S -> N.
FT VARIANT 193 193 N -> S.
FT VARIANT 301 301 I -> T.
FT VARIANT 318 318 K -> E.
FT VARIANT 403 403 Y -> N.
FT VARIANT 525 525 A -> R.
FT VARIANT 833 833 P -> L.
SQ SEQUENCE 848 AA; 96433 MM; B240019C8737BCB3 CRC64;

Query Match 89.1%; Score 2789.5; DB 15; Length 848;
Best Local Similarity 81.9%; Pred. No. 2.8e-230;
Matches 528; Conservative 22; Mismatches 28; Indels 67; Gaps 3;

QY 1 VEKMTVTVVYGVPPWKEATTTLLFCASDAKAYDVEVHNVMTACVPTDPNQEVLNVNT 60
DB 30 VEKMTVTVVYGVPPWKEATTTLLFCASDAKAYDVEVHNVMTACVPTDPNQEVLNVNT 89
QY 61 EHFNMKNKNMVEQOMEDIISLMDOSLKPVCVLTPLCGA----- 98
DB 90 EHFNMKNKNMVEQOMEDIISLMDOSLKPVCVLTPLCTVLTNCQDVANATTTSSSEGMERG 149
QY 99 -----GCDTSVITQACPISFE 115
DB 150 EIKNSFNITRSIRNKVQKEYALFYKLDVPIIDNSKNTKRYLLINCSTVITQACPVSFE 209
QY 116 PIPHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 175
DB 210 PIPHYCAPAGPAILKCKNNKTFNGKQCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 269
QY 176 IRSDFNNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 235
DB 270 IRSDFNNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 329
QY 236 ISRAKNDTLKQIVIKLRQFENKTIIVNHSFGDPEIVMSFNGGSEFFYCNSTQLFNS 295
DB 330 ISRAQNNNTLQIVIKLRQFENKTIIVFHSFGDPEIVMSFNGGSEFFYCNSTQLFNS 389
QY 296 TWNNNTESNNTGK-NTTLPCKRIKQIINMOEVGKAMYPPIRGQIRCSNITGLLNR 354
DB 390 TW-NDTEKSSGEGVDITILPCRIKQIINMOEVGKAMYPPIRGQIRCSNITGLLNR 448
QY 355 DGGINENGTEIFRPGGDMRDNRSELKYKVKVKTLEPGVATPKKRRVOREKAVGIG 414
DB 449 DGGKNESETEIFRPGGDMRDNRSELKYKVKVKTLEPGVATPKKRRVOREKAVGIG 508
QY 415 AVFLGFLGAAGSTMGASNTLTVQARLLLSGIVQOONNLRAIEAQOQMLQTLVGIQOL 474
DB 509 ALFLGFLGAAGSTMGASNTLTVQARQLLSGIVQOONNLRAIEAQOQMLQTLVGIQOL 568
QY 475 QARVLAVERYLGDQOLLGIWGSGLICTTAVPKNASNNKSLDPIINNNMTMEVERID 534
DB 569 QARVLAVERYLGDQOLLGIWGSGLICTTAVPKNASNNKSLDPIINNNMTMEVERID 628

```

```

QY 535 NYTSEIYTLIEESQNOEKNEOBLELDKWNASLNNWEDITNNLWY 579
DB 629 NYTWTIYTLIEESQIQOEKNEOBLELDKWNASLNNWEGITKWLWY 673

RESULT 6
ID 040222 PRELIMINARY; PRT; 854 AA.
AC 040222:
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RX MEDLINE=96432129; PubMed=8835195;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.;
RT "Construction and characterization of a stable full-length macrophage-
RT tropic HIV type 1 molecular clone that directs the production of high
RT titers of progeny virions.";
RL AIDS Res. Hum. Retroviruses 12:191-194(1996).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.C.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004394; AAB64170.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KV AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MM; 06C45E691036C12 CRC64;

Query Match 89.0%; Score 2784.5; DB 15; Length 854;
Best Local Similarity 82.0%; Pred. No. 7.7e-230;
Matches 533; Conservative 12; Mismatches 34; Indels 71; Gaps 6;

QY 1 VEKMTVTVVYGVPPWKEATTTLLFCASDAKAYDVEVHNVMTACVPTDPNQEVLNVNT 60
DB 30 VENLMTVTVVYGVPPWKEATTTLLFCASDAKAYDVEVHNVMTACVPTDPNQEVLNVNT 89
QY 61 EHFNMKNKNMVEQOMEDIISLMDOSLKPVCVLTPLCGA----- 98
DB 90 EHFNMKNKNMVEQOMEDIISLMDOSLKPVCVLTPLCTVLTNCQDVANATTTSSSEGMERG 149
QY 99 -----GCDTSVITQACPISFEPI 117
DB 150 IKNSFNITTSIRDKVKKDYALFYRLDVPIIDNNTSYRLINCSTVITQACPVSFEPI 209
QY 118 PIPHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENVIR 177
DB 210 PIPHYCAPAGPAILKCKKFFGTGTPCKNVSTVOCTHGIRPVSTQLLNGSLAEENVIR 269
QY 210 PIPHYCAPAGPAILKCKKFFGTGTPCKNVSTVOCTHGIRPVSTQLLNGSLAEENVIR 269
QY 178 SDNFNNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHNIS 237
DB 270 SSNTFDNNKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHNIS 329
QY 238 RAKNNDTLKQIVIKLRQF-ENKTIIVNHSFGDPEIVMSFNGGSEFFYCNSTQLFNST 296
DB 330 RTKNNNTLNLQATLKQFQGNKTIIVNHSFGDPEIVMSFNGGSEFFYCNSTQLFNST 389
QY 297 WNN-----NTESNNTGK-NTTLPCKRIKQIINMOEVGKAMYPPIRGQIRCSNITGL 351
DB 390 WNFNGTWNLTQSNQTEGNDITLPCRIKQIINMOEVGKAMYPPIRGQIRCSNITGLI 449

```

```

QY 352 LTRDGINENG-TEIPRPGGDMRDNRSELYKVKVYKIEPLGVAPTKCRKRVQREKRA 410
DB 450 LTRDGNHNNNDTEFPFGGDMRDNRSELYKVKVYKIEPLGVAPTKCRKRVQREKRA 509
QY 411 VG-ICAVFLGFLGAAGSTMGAASTMTLVQARLLLSGIVQOONNLLRAIEAOQRLQLTWG 469
DB 510 VGTICAMEFGFLGAAGSTMGAASTMTLVQARLLLSGIVQOONNLLRAIEAOQRLQLTWG 569
QY 470 GIKQOLARVLAVERLYGDOQLGIMGCGSKLICCTAVPWNASMSKSLDRIMNNNTMWE 529
DB 570 GIKQOLARVLAVERLYRDOQLGIMGCGSKLICCTAVPWNASMSKSLDRIMNNNTMWE 629
QY 530 EREIDNTSEIYTLIEESQNOEKNEQELBLDKKASLNNWPDITNMLWY 579
DB 630 EREIDNTGLIYTLIEESQNOEKNEQELBLDKKASLNNWPDITNMLWY 679

RESULT 7
092761 PRELIMINARY; PRT; 852 AA.
092761:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SFMS2.9;
RX MEDLINE=98178716; PubMed=9519894; Salminen M.O., Carr J.K.,
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337 (1998).
DR EMBL; AF025750; AAC40588.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000173; GAP_dhdtogase.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
DR PROSITE; PS00071; GARDP.1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB0611B617 CRC64;

Query Match 88.9%; Score 2780.5; DB 15; Length 852;
Best Local Similarity 81.1%; Pred. No. 1,7e-229;
Matches 525; Conservative 20; Mismatches 31; Indels 71; Gaps 4;
3 KLMVTYVYGVYVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNNTV 62
DB 32 QLMVTYVYGVYVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNNTV 91
QY 63 FNMKNWVQWQEDIIISLMDOSLKPCVLTPLCG----- 97
DB 92 FNMKNWVQWQEDIIISLMDOSLKPCVLTPLCG----- 151
QY 98 -----AGCDTSVITQACPKISPE 115
DB 152 IKNGSFNITTSIRDKQKDYALFYRLDVVPIDTNTNTSYRLVSCNTSITITQACPKISPE 211
QY 116 PIPIHYCAPAGFAILKCKNDKTFNGKQPCKNVSTVOCTHGIRPVYSTOLLNGLAEEVY 175
DB 212 PIPIHYCAPAGFAILKCKNDKTFNGKQPCKNVSTVOCTHGIRPVYSTOLLNGLAEEVY 271
QY 176 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTYTGIIIDIRQAHCN 235
DB 272 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTYTGIIIDIRQAHCN 331
QY 236 ISRAKMDTLKQIVIKLAEQFENKTIIVFNHSSGGDEIWMHSHFNGGGEFFYCNSTOLFNS 295

```

```

DB 332 ISRAKMDTLKQIVIKLAEQFENKTIIVFNHSSGGDEIWMHSHFNGGGEFFYCNSTOLFNS 391
QY 236 TMANNTBESNNTGEG--TITLPCRKOIINMVEGKAMYPPIGQIRCCSNTIGLLT 353
DB 332 TW-NDTBSNNTREDDTITLPCRKOIINMVEGKAMYPPIGQIRCCSNTIGLLT 450
QY 354 RDGINENG-TEIPRPGGDMRDNRSELYKVKVYKIEPLGVAPTKCRKRVQREKRAV 412
DB 451 RDGINNGTGTETFPFGGDMRDNRSELYKVKVYKIEPLGVAPTKCRKRVQREKRAV 510
QY 413 ICAVFLGFLGAAGSTMGAASTMTLVQARLLLSGIVQOONNLLRAIEAOQRLQLTWG 472
DB 511 ICAVFLGFLGAAGSTMGAASTMTLVQARLLLSGIVQOONNLLRAIEAOQRLQLTWG 570
QY 473 QLOARVLAVERLYGDOQLGIMGCGSKLICCTAVPWNASMSKSLDRIMNNNTMWE 532
DB 571 QLOARVLAVERLYRDOQLGIMGCGSKLICCTAVPWNASMSKSLDRIMNNNTMWE 630
QY 533 IDNTSEIYTLIEESQNOEKNEQELBLDKKASLNNWPDITNMLWY 579
DB 631 IDNTGLIYTLIEESQNOEKNEQELBLDKKASLNNWPDITNMLWY 677

RESULT 8
003811 PRELIMINARY; PRT; 853 AA.
003811:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Surface envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ADA;
RX MEDLINE=91195299; PubMed=2014229;
RA Westervelt P., Gendelman H.E., Ratner L.;
RT "Identification of a determinant within the human immunodeficiency
RT virus 1 surface envelope glycoprotein critical for productive
RT infection of primary monocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101 (1991).
DR EMBL; M60472; AAA5065.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
KW Envelope protein.
SQ SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

Query Match 88.8%; Score 2777; DB 15; Length 853;
Best Local Similarity 81.4%; Pred. No. 3.4e-229;
Matches 528; Conservative 18; Mismatches 33; Indels 70; Gaps 5;
1 VEKLMVTYVYGVYVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNNTV 60
DB 30 VENLMVTYVYGVYVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNNTV 89
QY 61 EHFNMKNWVQWQEDIIISLMDOSLKPCVLTPLCGA----- 98
DB 90 EHFNMKNWVQWQEDIIISLMDOSLKPCVLTPLCGA----- 149
QY 99 -----GCDTSVITQACPKISPE 117
DB 150 IKNGSFNITTSIRDKQKDYALFYRLDVVPIDNNTSYRLVSCNTSITITQACPKISPE 209
QY 118 PIPIHYCAPAGFAILKCKNDKTFNGKQPCKNVSTVOCTHGIRPVYSTOLLNGLAEEVY 177
DB 210 PIHYCTPPAGFAILKCKNDKTFNGKQPCKNVSTVOCTHGIRPVYSTOLLNGLAEEVY 269
QY 178 SDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTYTGIIIDIRQAHCN 237

```

```

Db 270 SNNFTDAKNIITVQKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCNIS 329
Qy 238 RAKNDITLQKIVLREOF-EKNTIVFNHSGSGPELVNHSFNGCGEPFYCNSTOLFNST 296
Db 330 RTKMNNTLNQIATLKQKQFNNKTIIVNQSOGDPETVMSFNGGEPFYCNSTOLFNST 389
Qy 297 WNN-----NTGSSNTEGN-TITLPCRIRKQIIMMOEKGKAYAPPIRGQIRCSSNITGL 351
Db 390 WNFNGTNNLOSNGTBNNDITLPCRIRKQIIMMOEKGKAYAPPIRGQIRCSSNITGL 449
Qy 352 LTRDGINENGTETFRPGGDMRDNRSELYKRYVKIIEPLGVAPTKRRRVQREKRAV 411
Db 450 LTRDGINSSGSEIFRPGGDMRDNRSELYKRYVKIIEPLGVAPTKRRRVQREKRAV 509
Qy 412 G-IGAVLFLGFGAAGSTMGAASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVMG 470
Db 510 GTTGAMFLGFGAAGSTMGAASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVMG 569
Qy 471 IKQARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 530
Db 570 IKQARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 629
Qy 531 REIDNTSEIYTLIEESQOQEKNEOBLELDKNASLNNMFDITNNLMY 579
Db 630 REIDNTSEIYTLIEESQOQEKNEOBLELDKNASLNNMFDITNNLMY 678

```

## RESULT 9

```

092762 ID 092762 PRELIMINARY; PRT; 854 AA.
AC 092762;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SFMS8.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
KW AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL: AF025754; AAC40591.1; -
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62CSB27D CRC64;

```

```

Query Match 88.5%; Score 2770; DB 15; Length 854;
Best Local Similarity 80.9%; Pred. No. 1.4e-228;
Matches 525; Conservative 23; Mismatches 29; Indels 72; Gaps 5;

```

```

Qy 2 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db 32 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 91
Qy 62 HFNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCGA----- 98
Db 92 NFNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCGA----- 151
Qy 99 -----GCDIVTTQACPCKISPE 115
Db 152 GEIKNGSFNITNRRGKMQKEVALFYKLDVVSIDNDNTSYRLISCSNTSVITQACPCKISFO 211

```

```

Qy 116 PIPHYCAPAGFALIKCNDKTENGKPCKNVSTVQCTHGIRPVYSTQLNLNGSLAEDEV 175
Db 212 PIPHYCAPAGFALIKCNDKTENGKPCKNVSTVQCTHGIRPVYSTQLNLNGSLAEDEV 271
Qy 176 IRSDFNFTNNAKTIIVQKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCN 235
Db 272 IRSDFNFTNNAKTIIVQKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCN 331
Qy 236 ISRAKNDITLQKIVLREOF-EKNTIVFNHSGSGPELVNHSFNGCGEPFYCNSTOLFNS 295
Db 332 ISRAKNDITLQKIVLREOF-EKNTIVFNHSGSGPELVNHSFNGCGEPFYCNSTOLFNS 391
Qy 296 WNN-----NTGSSNTEGN-TITLPCRIRKQIIMMOEKGKAYAPPIRGQIRCSSNITGL 353
Db 392 WNFNGTNNLOSNGTBNNDITLPCRIRKQIIMMOEKGKAYAPPIRGQIRCSSNITGL 449
Qy 354 RDGINENGT--EIRPGGDMRDNRSELYKRYVKIIEPLGVAPTKRRRVQREKRA 410
Db 452 RDGINENGT--EIRPGGDMRDNRSELYKRYVKIIEPLGVAPTKRRRVQREKRA 510
Qy 411 VGIGAVLFLGFGAAGSTMGAASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVMG 470
Db 511 VGIGAVLFLGFGAAGSTMGAASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVMG 570
Qy 471 IKQARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 530
Db 571 IKQARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 630
Qy 531 REIDNTSEIYTLIEESQOQEKNEOBLELDKNASLNNMFDITNNLMY 579
Db 631 REIDNTSEIYTLIEESQOQEKNEOBLELDKNASLNNMFDITNNLMY 679

```

## RESULT 10

```

056110 ID 056110 PRELIMINARY; PRT; 851 AA.
AC 056110;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SFMS8.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
KW AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL: AF025756; AAC40593.1; -
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDA833BE6464531 CRC64;

```

```

Query Match 88.5%; Score 2769; DB 15; Length 851;
Best Local Similarity 81.0%; Pred. No. 1.6e-228;
Matches 524; Conservative 22; Mismatches 31; Indels 70; Gaps 3;

```

```

Qy 2 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db 31 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy 62 HFNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCG----- 97

```



```

Db      91 NNNMMKNNNVEQMEDIISLMDQSLKPCVKLTPLCVTLNCLDRLNDNTNTSSSGEIMERG 150
Qy      98 -----ACGDSVITQACPKISFEPI 117
Db      151 EIKNSFNITTSIRDKVQKEYALLHLKLDVVPIDNDTSTRVLSGNSVITQACPKISFEPI 210
Qy      118 PIHYCAPAGFALLKCNKDTFNGKPCPKNVSTVOCTHGIRPVVSTOLLNGSLAESEVIR 177
Db      211 PIHYCAPAGFALLKCNKDRFNGKPCPKNVSTVOCTHGIRPVVSTOPLNGSLAESEVIR 270
Qy      178 SDFNNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRQACHNS 237
Db      271 SDFNNNAKTIIVOLNESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRQACHNS 330
Qy      238 RAKNMDTLKOIIVIKLREOFENKTIIVFNHSSGGDPEIWMHSFNGCGEFPYCNSTOLFNSW 297
Db      331 EAKMHTLBOJLAKKREOFENKTIIVFNHSSGGDPEIWMHSFNGCGEFPYCNSTOLFNSW 390
Qy      238 N-----NNTGSGNNTGNTITLPCRJKOIINMMQEVGKAMVAPIRGOIRCSNITGLL 352
Db      391 SVNSTWNTGEGVNNI-GENITLPCRJKOIINMMQEVGKAMVAPIRGOIRCSNITGLL 449
Qy      333 TRDGINENGTETFRPGGDMRDNMRSELYKVKVIEPLGVAPTKCRVYQREKRAVG 412
Db      450 TRDGSNNNETETFRPGGDMRDNMRSELYKVKVIEPLGVAPTKCRVYQREKRAVG 509
Qy      413 IGAVLGLFGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRLQITVWGK 472
Db      510 IGAVLGLFGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAOQRLQITVWGK 569
Qy      473 QLOARVLAVERYLGDQQLGIWCGSKLICCTAVPWNASWSNKSIDRIINNNTMMEWRE 532
Db      570 QLOARVLAVERYLKDQQLGIWCGSKLICCTAVPWNASWSNKSIDRIINNNTMMEWRE 629
Qy      533 IDNTSSEIYTLIESONOEKNEQELBELDKVASLNNPDTNMLWY 579
Db      630 INNTSLIYTLIESONOEKNEQELBELDKVASLNNPDTNMLWY 676

```

## RESULT 11

```

Q77989 PRELIMINARY; PRT; 863 AA.
ID Q77989;
AC Q77989;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
Envelope glycoprotein.
ENV.
Human immunodeficiency virus 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBT_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SF128A;
MEDLINE=91056585; PubMed=2243391;
RA Liu Z.-Q., Wood C., Levy J.A., Cheng-Mayer C.;
"the viral envelope gene is involved in macrophage tropism of a human immunodeficiency virus type 1 strain isolated from brain tissue.";
RL J. Virol. 64:6148-6153 (1990).
DR EMBL; M95292; AAA44331.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SEQUENCE 863 AA; 98222 MW; 0157397FA8B4C5D7 CRC64;

```

```

Query Match      88.3%; Score 2764; DB 15; Length 863;
Best Local Similarity 79.4%; Pred. No. 4, 5e-228;
Matches 523; Conservative 20; Mismatches 34; Indels 82; Gaps 3;
Qy      2 EKLMTVYVYGVPWKEATTLTFCASDAKAYDTEVHNVNATHACVPTDPPQEVLENVTE 61

```

```

Db      31 EKLMTVYVYGVPWKEATTLTFCASDAKAYDTEVHNVNATHACVPTDPPQEVLENVTE 90
Qy      62 HNNMMKNNNVEQMEDIISLMDQSLKPCVKLTPLCGA----- 98
Db      91 NNNMMKNNNVEQMEDIISLMDQSLKPCVKLTPLCVTLNCLDRLNDNTNTDNTNATSS 150
Qy      99 -----GCDTSVITQAC 109
Db      151 SLRGTEGIIKNSFNITTSIRDKVQKEYALFYKLDIVPIINNNTYRLINCTSTITQAC 210
Qy      110 PKISFEPIPIHYCAPAGFALLKCNKDTFNGKPCPKNVSTVOCTHGIRPVVSTOLLNGSL 169
Db      211 PKVSEPIPIHYCTPAGFALLKCNKDRFNGKPCPKNVSTVOCTHGIRPVVSTOLLNGSL 270
Qy      170 AEEVVISSDNTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDI 229
Db      271 AEEVVISSDNTNNAKTIIVOLNESVVINCTRPNNNTRKSHIGPRAFYTTGIIIGDI 330
Qy      230 ROAHGNISRAKNDTLKOIIVIKLREOFENKTIIVFNHSSGGDPEIWMHSFNGCGEFPYCN 289
Db      331 ROAHCTLKTQWMDTLROIALKREOFENKTIIVFNHSSGGDPEIWMHSFNGCGEFPYCN 390
Qy      230 TOLFNSWN-----NNTGSGNNTGNTITLPCRJKOIINMMQEVGKAMVAPIRGO 340
Db      391 TRLFNSTNDTRLFNSTWNTTIGSNTGNT-ITLPCRJKOIINMMQEVGKAMVAPIRGO 449
Qy      341 TRCSNITGLLTLTRDGINENGTETFRPGGDMRDNMRSELYKVKVIEPLGVAPTKCR 400
Db      450 TRCSNITGLLTLTRDGINENGTETFRPGGDMRDNMRSELYKVKVIEPLGVAPTKCR 509
Qy      401 RRVOREKRAVGIGAVFLGFGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAO 460
Db      510 RRVOREKRAVGIGAVFLGFGAAGSTMGAASMTLTVQARLLSGIVQOONNLLRAIEAO 569
Qy      461 QRLQITVWGKIQLOARVLAVERYLGDQQLGIWCGSKLICCTAVPWNASWSNKSIDRI 520
Db      570 QRLQITVWGKIQLOARVLAVERYLGDQQLGIWCGSKLICCTAVPWNASWSNKSIDRI 629
Qy      521 WNNMTMMEBERIDNTSSEIYTLIESONOEKNEQELBELDKVASLNNPDTNMLWY 579
Db      630 WNNMTMMEBERIDNTSSEIYTLIESONOEKNEQELBELDKVASLNNPDTNMLWY 688

```

## RESULT 12

```

O56562 PRELIMINARY; PRT; 851 AA.
ID O56562;
AC O56562;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
Envelope glycoprotein.
ENV.
Human immunodeficiency virus 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBT_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PH1159;
MEDLINE=99372987; PubMed=10445815;
RA Ataman-Onal Y., Coiffier C., Girard A., Babić-Bregg A., Biron F.,
Verrier B.;
"Comparison of complete env gene sequences from individuals with symptomatic primary HIV type 1 infection.";
RL AIDS Res. Hum. Retroviruses 15:1035-1039 (1999).
DR EMBL; AF041128; AAC02519.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SEQUENCE 851 AA; 96811 MW; 519363B6FB8BF146 CRC64;

```

Query Match 88.2%; Score 2759; DB 15; Length 851;  
 Best Local Similarity 80.5%; Pred. No. 1.2e-227;  
 Matches 520; Conservative 24; Mismatches 34; Indels 68; Gaps 4;

QY 2 EKLWVTVYGVVWPKKEATTTLFCASDAKAYDEVHNMATHACVPDPNPOEVLNENVT 61  
 DB 31 EKLWVTVYGVVWPKKEATTTLFCASDAKAYDEVHNMATHACVPDPNPOEVLNENVT 90  
 QY 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 98  
 DB 91 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGVLTPLCTDLDGNATNTMKEGKNCSE 150  
 QY 99 -----GCDTSVITQACPKISFPIPIHYCA 123  
 DB 151 NYTTIRDKYQKAYALFYKLDVQWQMEDNTSYRLISCTSVITQACPKISFPIPIHYCA 210  
 QY 124 PAGFAILKNDKTNGKPGCKNVSTVQCTHGIRPVYSTQLLNGSLAEEVVISDNFTN 183  
 DB 211 PAGFAILKNDKTNGKPGCKNVSTVQCTHGIRPVYSTQLLNGSLAEEVVISDNFTN 270  
 QY 184 NAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNISRAKMD 243  
 DB 271 NTKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNISRAKMD 330  
 QY 244 TLKQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFNSTWNN-- 300  
 DB 331 TLNQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFNSTWNN-- 390  
 QY 301 --TGSNNTEG--NTITLPCRIKQIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 356  
 DB 391 NVTEGSRRTTGGNDTLIPCRKQIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 450  
 QY 357 GIN---ENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKRRVQREKRAVGI 413  
 DB 451 GNNNENENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKRRVQREKRAVGI 510  
 QY 414 GAVFLGLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQMLDLTWGIKQ 473  
 DB 511 GAVFLGLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQMLDLTWGIKQ 570  
 QY 474 LQAVLAVERLYGDOOLIGIWCSSGKLICTAVPNNASWSKSLDRIMNMTMWEEREI 533  
 DB 571 LQAVLAVERLYGDOOLIGIWCSSGKLICTAVPNNASWSKSLDRIMNMTMWEEREI 630  
 QY 534 DNTSEIYTLIEESQNOQEKNEOELLELDKWSLNNWFDITNMLWY 579  
 DB 631 DNTSEIYTLIEESQNOQEKNEOELLELDKWSLNNWFDITNMLWY 676

RESULT 13  
 Q9PKW7 PRELIMINARY; PRT; 846 AA.  
 AC Q9PKW7; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE GP160, envelope glycoprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9189881; PubMed=8446773;  
 RA Schmittgeneyerova H., Gayet O., Guettari N., Bolmont C., Hirsch I.,  
 RT "Characterization of HIV1-PAR, a macrophage-tropic strain: cell  
 RT tropism, virus/cell entry and nucleotide sequence of the envelope  
 RT glycoprotein."  
 RL Res. Virol. 144:21-26(1993).  
 DR InterPro: IPR000328; ENV\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

Query Match 88.0%; Score 2752.5; DB 15; Length 846;  
 Best Local Similarity 81.1%; Pred. No. 4.4e-227;  
 Matches 522; Conservative 23; Mismatches 30; Indels 69; Gaps 5;

QY 2 EKLWVTVYGVVWPKKEATTTLFCASDAKAYDEVHNMATHACVPDPNPOEVLNENVT 61  
 DB 31 EKLWVTVYGVVWPKKEATTTLFCASDAKAYDEVHNMATHACVPDPNPOEVLNENVT 90  
 QY 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 98  
 DB 91 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGVLTPLCTDLDGNATNTMKEGKNCSE 150  
 QY 99 -----GCDTSVITQACPKISFPIPIHYCA 116  
 DB 151 EIKNSFNITTSIRDKYQKAYALFYKLDVVPIDNSTRYRLISCTSVITQACPKVTEP 210  
 QY 117 IPIHYCAPAGFAILKNDKTNGKPGCKNVSTVQCTHGIRPVYSTQLLNGSLAEEVVI 176  
 DB 211 IPIHYCAPAGFAILKNDKTNGKPGCKNVSTVQCTHGIRPVYSTQLLNGSLAEEVVI 270  
 QY 177 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNT 236  
 DB 271 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNT 330  
 QY 237 SRAKMDTLKQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFNST 296  
 DB 331 SKTKMETFPQIVKRLKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFNST 390  
 QY 297 MNNTEGSSNTEGNTITLPCRIKQIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 356  
 DB 391 W-NDTESSNTE-RTITLPCRIKQIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 448  
 QY 357 G-INENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKRRVQREKRAVGI 415  
 DB 449 GNNNNGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKRRVQREKRAVGI 508  
 QY 416 VFGLGLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQMLDLTWGIKQ 475  
 DB 509 VFGLGLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQMLDLTWGIKQ 568  
 QY 476 ARVLAVERLYGDOOLIGIWCSSGKLICTAVPNNASWSKSLDRIMNMTMWEEREIDN 535  
 DB 569 ARVLAVERLYGDOOLIGIWCSSGKLICTAVPNNASWSKSLDRIMNMTMWEEREIDN 627  
 QY 536 YTSSEIYTLIEESQNOQEKNEOELLELDKWSLNNWFDITNMLWY 579  
 DB 628 YTSSEIYTLIEESQNOQEKNEOELLELDKWSLNNWFDITNMLWY 671

RESULT 14  
 Q8AR21 PRELIMINARY; PRT; 847 AA.  
 AC Q8AR21; 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Envelope glycoprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=SHIVSF162P3.2;  
 RA Gao F.;  
 RL Submitted (JUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF536757; AAN05642.1; --  
 DR SEQUENCE 847 AA; 95879 MW; 9CF863B97B3BB54C CRC64;

Query Match 88.0%; Score 2752; DB 15; Length 847;  
 Best Local Similarity 81.4%; Pred. No. 4.7e-227;  
 Matches 525; Conservative 17; Mismatches 35; Indels 68; Gaps 4;

QY 1 VKLWTVVYGVVWKEATTTTFCASDAKAYDTEVANNWATHACVTPDNPQEVLENT 60  
 |||  
 DB 30 VKLWTVVYGVVWKEATTTTFCASDAKAYDTEVANNWATHACVTPDNPQEVLENT 89  
 |||

QY 61 HFNMMKNNVBOEMODIISLMDQSLKPCVKLTPLCGA----- 98  
 |||  
 DB 90 HFNMMKNNVBOEMODIISLMDQSLKPCVKLTPLCVTLCTNLNENATNTSSNMKNR 149  
 |||

QY 99 -----GCDTSVITQACPISPE 115  
 |||  
 DB 150 GEIKNCSFVNTTSIGNKQKEYALFYRLDVPIDNDNTSYNLINCTSVITQACPVSPE 209  
 |||

QY 116 PIPHYCAPAGFAILKCKNDKTRNGKPCXNVSTVOCTHGIRPVSTOLLNGSLAEEVY 175  
 |||  
 DB 210 PIPHYCAPAGFAILKCKNDKTRNGKPCXNVSTVOCTHGIRPVSTOLLNGSLAEEVY 269  
 |||

QY 176 IISDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRAHCN 235  
 |||  
 DB 270 IISDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRAHCN 329  
 |||

QY 236 ISRAKNDTLKQIVIKLREOFENKTIIVNHSQGDPEIVMHSFNCGEFFYCNSTOLFNS 295  
 |||  
 DB 330 ISRAKNDTLKQIVIKLREOFENKTIIVNHSQGDPEIVMHSFNCGEFFYCNSTOLFNS 389  
 |||

QY 236 TNNNTGEGNNTGNTTTPCRKIQIINMOEYKAMVAPPIRQIRGSSNTGLLTLD 355  
 |||  
 DB 390 TNNNTGEGNNTGNTTTPCRKIQIINMOEYKAMVAPPIRQIRGSSNTGLLTLD 447  
 |||

QY 356 GGIN-ENGTEIFRPGGDMRDMRSELKYKVKIIEPLGVAIPKCRARVOREKAVGIG 414  
 |||  
 DB 448 GGIN-ENGTEIFRPGGDMRDMRSELKYKVKIIEPLGVAIPKCRARVOREKAVGIG 507  
 |||

QY 415 AVFLFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLIRAIIEAQRMLOLTWNGIKOL 474  
 |||  
 DB 508 AVFLFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLIRAIIEAQRMLOLTWNGIKOL 567  
 |||

QY 475 QARVLAVERYLQDOQLIGWCGSGKICTTAVPWNASNSKSLDRIINNMWEMEREID 534  
 |||  
 DB 568 QARVLAVERYLQDOQLIGWCGSGKICTTAVPWNASNSKSLDRIINNMWEMEREID 627  
 |||

QY 535 NTSRIYTLIESONOQEKNEOELLELDKASISLMMNFDTITNKLWY 579  
 |||  
 DB 628 NTSRIYTLIESONOQEKNEOELLELDKASISLMMNFDTITNKLWY 672  
 |||

QY 070150 PRELIMINARY; PRT; 843 AA.  
 070150  
 AC 070150, 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=014;  
 RA MEDLINE=95194694; PubMed=7888189;  
 RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,  
 RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.,  
 RT "Genetic variation of HIV type 1 in four World Health Organization-  
 RT sponsored vaccine evaluation sites: generation of functional envelope  
 RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,  
 RT and E. WHO Network for HIV Isolation and Characterization.";  
 RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).  
 RN (2)

RP SEQUENCE FROM N.A.  
 RC STRAIN=014;  
 RX MEDLINE=95194694; PubMed=7545977;  
 RA WHO Global Programme on AIDS;  
 RT "HIV type 1 variation in World Health Organization-sponsored vaccine  
 RT evaluation sites: genetic screening, sequence analysis, and  
 RT preliminary biological characterization of selected viral strains. WHO  
 RT Network for HIV Isolation and Characterization.";  
 RL AIDS Res. Hum. Retroviruses 10:1327-1343(1994).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=014;  
 RX MEDLINE=96190564; PubMed=8627686;  
 RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,  
 RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,  
 RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,  
 RA Hahn B.H.,  
 RT "Molecular cloning and analysis of functional envelope genes from  
 RT human immunodeficiency virus type 1 sequence subtypes A through G. The  
 RT WHO and NIAID Networks for HIV Isolation and Characterization.";  
 RL J. Virol. 70:1651-1657(1996).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=014;  
 RA Allen E.E.;  
 RL Submitted (Apr-1994) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; U08801; AAB05185.1; -  
 DR InterPro; IPR000328; Env\_GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
 FT NON-TER  
 SQ SEQUENCE 843 AA; 95679 MW; 14DB49A6D2F5CE4 CRC64;

Query Match 87.8%; Score 2747; DB 15; Length 843;  
 Best Local Similarity 80.9%; Pred. No. 1.2e-226;  
 Matches 518; Conservative 23; Mismatches 37; Indels 62; Gaps 3;

QY 2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVANNWATHACVTPDNPQEVLENT 61  
 |||  
 DB 29 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVANNWATHACVTPDNPQEVLENT 88  
 |||

QY 62 HFNMMKNNVBOEMODIISLMDQSLKPCVKLTPLCGA----- 98  
 |||  
 DB 89 HFNMMKNNVBOEMODIISLMDQSLKPCVKLTPLCVTLCTNLNENATNTSSNMGEIKN 148  
 |||

QY 99 -----GCDTSVITQACPISPEPIHY 121  
 |||  
 DB 149 CSFNITTSIKTKVKQYALFYKLDVVPIDNDNTSYNLINCNISVITQACPVSPEPIHY 208  
 |||

QY 122 CAPAGFAILKCKNDKTRNGKPCXNVSTVOCTHGIRPVSTOLLNGSLAEEVIRSDNF 181  
 |||  
 DB 209 CAPAGFAILKCKNDKTRNGKPCXNVSTVOCTHGIRPVSTOLLNGSLAEEVIRSDNF 268  
 |||

QY 182 TNNACTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRAHCNISRAKW 241  
 |||  
 DB 269 TNNACTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGIIIGDIRAHCNISRAKW 328  
 |||

QY 242 NDTLQIYIKLREOFENKTIIVNHSQGDPEIVMHSFNCGEFFYCNSTOLFNSTWNNT 301  
 |||  
 DB 329 NDTLQIYIKLREOFENKTIIVNHSQGDPEIVMHSFNCGEFFYCNSTOLFNSTWNNT 388  
 |||

QY 302 EGSNNTGNTTTPCRKIQIINMOEYKAMVAPPIRQIRGSSNTGLLTTRDGCINEN 361  
 |||  
 DB 389 EGSNNTGNTTTPCRKIQIINMOEYKAMVAPPIRQIRGSSNTGLLTTRDGCINEN 448  
 |||

QY 382 -GTEIFRPGGDMRDMRSELKYKVKIIEPLGVAIPKCRARVOREKAVGIG-IGAVFLG 419  
 |||  
 DB 449 -GTEIFRPGGDMRDMRSELKYKVKIIEPLGVAIPKCRARVOREKAVGIG-IGAVFLG 508  
 |||

QY 420 FLGAAGSTMGASMTLTVQARLLSGIVQOQNNLIRAIIEAQRMLOLTWNGIKOLQARVL 479  
 |||

Db	509	FLGASSTMGASITLTVQARQLLSGIVQOQRNLLRAIEAQOHLJOLTWGIKOLQARVL	568
Qy	480	AVERYLDGQOLLGIMGSGKLICTTAVPWNASWNSKSLDRIWNNMTWMEWEREIDNYTSE	539
Db	569	AVERYLDGQOLLGIMGSGKLICTTAVPWNASWNSKSLDKIWNNTWMEWEREIDNYTSE	628
Qy	540	IYTLIESQNOQEKNEOELLELDKWKASIMNWFDTNNLWY	579
Db	629	IYTLIESQNOQEKNEELLELDKWKASIMNWFDTNNLWY	668

Search completed: December 12, 2003, 12:32:39  
 Job time : 32.7807 secs



```

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL)
SEQUENCE 848 AA; 96475 MW; 2076751227EC3F3 CRC64;

Query Match 89.5%; Score 2799.5; DB 1; Length 848;
Best Local Similarity 82.2%; Pred. No. 7.1e-219;
Matches 530; Conservative 20; Mismatches 28; Indels 67; Gaps 3;

QY 1 VEKLMVTVYGVVPWKEATTTTFCASDAKAYDTEVHNVATTHACVPTDPNPOEVLEBNT 60
DB 30 VEKLMVTVYGVVPWKEATTTTFCASDAKADTEVHNVATTHACVPTDPNPOEVLEBNT 89
QY 61 EHFNMKNKNWQEQEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 90 EDFNMKNKNWQEQEDVINLMDQSLKPCVKLTPLCVTLNCKDVNATNTSSSEGMERG 149
QY 99 -----GQDSVTVQACPKISFE 115
DB 150 EIKNCSFNITSIIDKQKEXALFYKLDVPIIDKNKTKYRLISGNTSVITVQACPKISFE 209
QY 116 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLAAEEBVT 175
DB 210 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLAAEEBVT 269
QY 176 IRSDFNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQAHN 235
DB 270 IRSDFNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQAHN 329
QY 236 ISRAKNDTLKOIYKLEQEFNKTIIVNHSGGPPEIWMHSFNGGFFPCNSQTOLNS 295
DB 330 ISRAQNMNTLKOIYKLEQEFNKTIIVTHSSGGPPEIWMHSFNGGFFPCNSQTOLNS 389
QY 296 TWNNNTSGSNNTGEG-TITLPCRIKOIIMMOEVKAYAPPIRGOICSSNITGLLTLR 354
DB 390 TW-NDTESSSGTSGNDTIIIPCRIOIIMMOEVKAYAPPIRGOICSSNITGLLTLR 448
QY 355 DGGINENGEIIFRPGGDMRDWRSELYKYVVKIIEPLGVAFTKCRVOREKRAVIG 414
DB 449 DGGKNESEIIFRPGGDMRDWRSELYKYVVKIIEPLGVAFTKCRVOREKRAVIG 508
QY 415 AVFLGFLGAAGSTMAASMTLTVOARLLLSGIVQOONNLRAIEQOQMLQUTWIGIQOL 474
DB 509 ALFLGFLGAAGSTMAASMTLTVOARLLLSGIVQOONNLRAIEQOQMLQUTWIGIQOL 568
QY 475 QARVLAVERVYGDQDLGIWGGSGKLICTTAVPAMNASNKSLLDIIMNMNTMEWEREID 534
DB 569 QARVLAVERVYKDQDLGIWGGSGKLICTTAVPAMNTSSNSLSDIIMNMNTMEWEREID 628
QY 535 NYTSEIYTLIESQOQEKNEQELLELDKMASLMMWFDTTWLWY 579
DB 629 NYTNTIYTLIESQOQEKNEQELLELDKMASLMMWFDTTWLWY 673

```

```

AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11691;
OX
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M65024; AAA5072.1; -.
CC PDB: 10B5; 15-MAY-97.
CC HIV; M38428; ENV$SF162.
CC InterPro: IPR000328; Env.GP41.
CC InterPro: IPR000777; GP120.
CC Pfam: PF00516; GP120; 1.
CC DR Pfam: PF00517; GP41; 1.
CC KW AIDS; Coat protein; polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 53 73 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 118 203 BY SIMILARITY.
FT DISULFID 125 194 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

```

Query Match 88.3%; Score 2764; DB 1; Length 847;  
 Best Local Similarity 81.7%; Pred. No. 5.3e-216;  
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;

1 VEKLVWTVYGVVWPKAEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60  
 30 VEKLVWTVYGVVWPKAEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89  
 61 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCGA----- 98  
 90 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCGA----- 149  
 99 -----GCPTSVITQACPKISPE 115  
 150 GBTKMCKFVTTISIRNKQKEVALFYKLDVPIINDNTSYKLNCTNITSYITQACPKVSPE 209  
 116 PIPHYCAPAGAPAILKCNKDTNGKPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 175  
 210 PIPHYCAPAGAPAILKCNKDKKFNSSGPTNVSITVQCTHGRPVSTOLLNGLAEEVY 269  
 176 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNRKSHIGRGAFTTGGIIGDIAQAHN 235  
 270 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNRKSHIGRGAFTTGGIIGDIAQAHN 329  
 236 ISRAKMDTLKQIVIKLREPFENKTIIVFNHSGGPEIYMHSGGPEFYCNSTOLFNS 295  
 330 ISGKMNNTLKOIVTKLQAFQNKTIIVFQSSGSGPEIYMHSGGPEFYCNSTOLFNS 389  
 296 TWNNNTBESNTEGNTTLPCKIKOINNMQEVGKAMAPIRGOIRCSNNTGILLTRD 355  
 390 TW-NNTIGPNNNTG--TITLPCRIKOINRMQEVGKAMAPIRGOIRCSNNTGILLTRD 447  
 356 GGIN-ENGELPRPGGDMRDMRSELYKYKVKLEPGVATKCKRPRVQEKRAVGLG 414  
 448 GKEKISNTTEIFRPGGDMRDMRSELYKYKVKLEPGVATKCKRPRVQEKRAVGLG 507  
 415 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQRMQLTGWGIKOL 474  
 508 AMFLGFLGAGSTMGARSLTLTVQARLLSGIVQOONNLRAIEAQRMQLTGWGIKOL 567  
 475 QARVLAVERRYLKDQOLLGIGWCGSKLICCTAVPNNASMSKSLDIRINNMWEMEREID 534  
 568 QARVLAVERRYLKDQOLLGIGWCGSKLICCTAVPNNASMSKSLDIRINNMWEMEREID 627  
 535 NYSRIYTLIESQNOQEKNEBELDKMASLWMPDTTMMY 579  
 628 NITNLIYTLIESQNOQEKNEBELDKMASLWMPDISKWLWY 672

RESULT 3  
 ENV\_HVISC STANDARD; PRT; 856 AA.  
 AC P05878;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN Human immunodeficiency virus type 1 (SC isolate) (HIV-1).  
 OS Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11702;  
 OX NCBI\_TaxID=11702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88219542; PubMed=3369091;  
 RA Guigo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,  
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.,  
 RT "Envelope sequences of two new United States HIV-1 isolates."  
 RL Virology 164:531-536(1988).  
 CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN  
 CC 1984 IN SOUTHERN CALIFORNIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, M17450, -; NOT\_ANNOTATED\_CDS.  
 DR HIV, M17450, ENV55C.  
 DR InterPro, IPR000328, Env GP41.  
 DR InterPro, IPR000777, GP120.  
 DR Pfam, PF00516, GP120, 1.  
 DR Pfam, PF00517, GP41, 1.  
 KM AIDS, Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KM Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 510  
 FT CHAIN 511 856  
 FT SITE 760 760  
 FT DISULFID 53 73  
 FT DISULFID 118 206  
 FT DISULFID 125 197  
 FT DISULFID 130 160  
 FT DISULFID 219 247  
 FT DISULFID 228 239  
 FT DISULFID 296 330  
 FT DISULFID 376 439  
 FT DISULFID 383 412  
 FT CARBOHYD 87 87  
 FT CARBOHYD 129 129  
 FT CARBOHYD 135 135  
 FT CARBOHYD 140 140  
 FT CARBOHYD 143 143  
 FT CARBOHYD 159 159  
 FT CARBOHYD 163 163  
 FT CARBOHYD 189 189  
 FT CARBOHYD 198 198  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 302 302  
 FT CARBOHYD 331 331  
 FT CARBOHYD 338 338  
 FT CARBOHYD 354 354  
 FT CARBOHYD 360 360  
 FT CARBOHYD 384 384  
 FT CARBOHYD 394 394  
 FT CARBOHYD 400 400  
 FT CARBOHYD 405 405  
 FT CARBOHYD 442 442  
 FT CARBOHYD 457 457  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 625 625  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CMC64;

Query Match 88.0%; Score 2754.5; DB 1; Length 856;  
 Best Local Similarity 80.1%; Pred. No. 3.2e-215;  
 Matches 523; Conservative 26; Mismatches 27; Indels 77; Gaps 6;

2 EKLWTVYGVVWPKAEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 61  
 31 EKLWTVYGVVWPKAEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 90  
 62 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCGA----- 98

```

Db 91 NFNMMKNMNVQWQMHEDIISLMDQSLKPCVKLTPLCVLTNLTNLTNTNATNTSSNRG 150
Qy 99 -----GCDSVTITQACPKTS 113
Db 151 KMEGECMNCSPNITTSIRSVQKEVALFYKLDVVPIDNTSYTLTNCSTSYITQACPKVS 210
Qy 114 PEPPIPIHYCAPAGFAILKNDKTFNGKPCNKVSTVOCTHGIRPVSTOLLNGLAEE 173
Db 211 PEPPIPIHYCA-RWFAILNCKNKKEFGTGPCNTVSTVOCTHGIRPVSTOLLNGLAEE 269
Qy 174 VVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRQAH 233
Db 270 VVLRSENTDPAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRQAH 329
Qy 234 CNIRAKKNDTLKQIVILKROPEKNTIVFNHSSGSGDEIVMHSNCGGEPFYCNSTOLF 293
Db 330 CNIRAKKNDTLKQIVILKROPEKNTIVFNHSSGSGDEIVMHSNCGGEPFYCNSTOLF 389
Qy 294 NSTWNNNTGSGNTEGN-TITLPCRIKQIIMMOWEVRKAWAPPIRGQIRCSNITGILL 352
Db 390 SSTW-NGTEGSGNNTGNTITLPCRIKQIIMMOWEVRKAWAPPIRGQIRCSNITGILL 448
Qy 353 TRDGGINENG-----TEIFRPGGDMRDNRSELYKYKVKIIEPLGVAPTYAKRRVQRE 407
Db 449 TRDGGNSKNGSKNTEIFRPGGDMRDNRSELYKYKVKIIEPLGVAPTYAKRRVQRE 508
Qy 408 KRAVG-IGAVPLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQMLQ 466
Db 509 KRAVGTIGAMFLGFLGAGSTMGATSMTLTVQARLLSGIVQOQNNLRAIEAOQMLQ 568
Qy 467 TWMGIKQOLARVLAVERYLQDOQLIGWCSGKLICTTAVPMNASWNSKSLDRIMNNMTW 526
Db 569 TWMGIKQOLARVLAVERYLQDOQLIGWCSGKLICTTAVPMNASWNSKSLDRIMNNMTW 628
Qy 527 MEMEREIDNTYSEIYTLIEESQOQKNOBOLLEIDKASIMNFDTNNWLMY 579
Db 629 MEMEREIDNTYSLIYTLIEESQOQKNOBOLLEIDKASIMNFDTNNWLMY 681

RESULT 4
ENV_HV1BR STANDARD: PRT; 861 AA.
ID ENV_HV1BR STANDARD: PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
ENY.
OC Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retrovirus; Retroviridae; Lentivirus.
OC NCBI_TaxID=11686;
RX MEDLINE=8509333; PubMed=2981635;
RA Main-Hobson S., Sonigo P., Danos O., Cole S., Allison M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02013; AAB59751.1; -
DR EMBL; A04321; CA00352.1; -
DR PIR; A03975; VCLJLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV5BRU.
DR InterPro; IPR000328; Env_GP41.

```

```

DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1
FT CHAIN 31
FT CHAIN 516
FT DISULFID 517
FT DISULFID 54
FT DISULFID 119
FT DISULFID 126
FT DISULFID 131
FT DISULFID 131
FT DISULFID 223
FT DISULFID 233
FT DISULFID 301
FT DISULFID 383
FT DISULFID 390
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 146
FT CARBOHYD 161
FT CARBOHYD 165
FT CARBOHYD 191
FT CARBOHYD 191
FT CARBOHYD 202
FT CARBOHYD 235
FT CARBOHYD 239
FT CARBOHYD 246
FT CARBOHYD 267
FT CARBOHYD 281
FT CARBOHYD 294
FT CARBOHYD 300
FT CARBOHYD 306
FT CARBOHYD 337
FT CARBOHYD 344
FT CARBOHYD 361
FT CARBOHYD 391
FT CARBOHYD 397
FT CARBOHYD 402
FT CARBOHYD 411
FT CARBOHYD 453
FT CARBOHYD 468
FT CARBOHYD 616
FT CARBOHYD 621
FT CARBOHYD 630
FT CARBOHYD 642
FT CARBOHYD 679
FT CARBOHYD 755
FT CARBOHYD 821
SQ SEQUENCE 861 AA; 97487 MW; 04DE2B4D4E4FD63A CRC64;

Query Match 87.1%; Score 2724.5; DB 1; Length 861;
Best Local Similarity 79.6%; Pred. No. 8.6e-213;
Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

Qy 2 EKLWTVYVYGVVVKKATTTTFCASDAKAYDTEVHNWMAFHACVPTDPNPOEVLNTE 61
Db 32 EKLWTVYVYGVVVKKATTTTFCASDAKAYDTEVHNWMAFHACVPTDPNPOEVLNTE 91
Qy 62 HFNMMKNMNVQWQMHEDIISLMDQSLKPCVKLTPLCGA----- 98
Db 92 NFNMMKNMNVQWQMHEDIISLMDQSLKPCVKLTPLCVLTNLTNLTNTNATNTSSNRG 151
Qy 99 -----GCDSVTITQACPKTS 113
Db 152 MMEKEGIKCSFNISTIRSVQKEVALFYKLDIIPIDNTSYTLTNCSTSYITQACPKVS 211
Qy 111 KISPEPIPIHYCAPAGFAILKNDKTFNGKPCNKVSTVOCTHGIRPVSTOLLNGLA 170
Db 212 KISPEPIPIHYCAPAGFAILKNDKTFNGKPCNKVSTVOCTHGIRPVSTOLLNGLA 271
Qy 171 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRQAH 228

```



Db 272 EEEVIRSANFTDNKATIIIVOLNQSVEINCTRPNNTRKSIIRIQRGPRAPVTTGK-IGN 330

Qy 229 IROACHNISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSSGGPPEIWMSPNCGEPFYC 287

Db 331 MQOACHNISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSSGGPPEIWMSPNCGEPFYC 390

Qy 288 NSTOLFNSWTWN--NTEGSNNTG-NTTLPCRIKQIIMNOVEGKAMYPAPPIRQIQC 343

Db 391 NSTOLFNSWTWNSTEGSNTGSDTLTLCRIQIIMNOVEGKAMYPAPPIRQIQC 450

Qy 344 SNNITGLLTRDGGINENGTETFRPGGDMRDNRSELKYKVKVETLEPGLVAPTKCKRRV 403

Db 451 SNNITGLLTRDGGNNNSSEIFRPGGDMRDNRSELKYKVKVETLEPGLVAPTKCKRRV 510

Qy 404 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVQOONNLRAIEAQOM 463

Db 511 VOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVQOONNLRAIEAQOM 570

Qy 464 IOLTWGIGIKQIARVAVRYLGDQQLGIGWCGSKLICCTAVPNWASNSKSLDRIMNN 523

Db 571 IOLTWGIGIKQIARVAVRYLGDQQLGIGWCGSKLICCTAVPNWASNSKSLDRIMNN 630

Qy 524 MTWMEWERINDYTSIYTLIESQOQKNEOELLELDKASLWNPETITWLY 579

Db 631 MTWMEWERINDYTSIYTLIESQOQKNEOELLELDKASLWNPETITWLY 686

ENV\_HV1J3 STANDARD; PRT; 867 AA.

AC P12489; 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last annotation update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].

GN Env.

OS Human immunodeficiency virus type 1 (HIV-1), isolate (HIV-1),

OC Viruses; Retroviridae; Retroviridae; Lentivirus.

CC NCBI\_Taxid=11694;

RP SEQUENCE FROM N.A.

RA MEDLINE=89352108; Pubmed=2669897;

RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;

RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria."

RL AIDS Res. Hum. Retroviruses 5:411-419(1989).

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

-----

CC EMBL\_M21138; AAB03526.1; -

DR HIV\_M21138; ENVSJH3.

DR InterPro; IPR000328; Env\_GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.

FT SIGNAL 1 30

FT CHAIN 31 516

FT CHAIN 517 867

FT DISULFID 53 73

FT DISULFID 118 217

FT DISULFID 125 208

FT DISULFID 130 160

FT DISULFID 230 259

FT DISULFID 240 251

EXTERIOR MEMBRANE GLYCOPROTEIN.

TRANSMEMBRANE GLYCOPROTEIN.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

FT DISULFID 308 342

FT DISULFID 388 437

FT DISULFID 395 450

FT CARBOHYD 87 87

FT CARBOHYD 135 135

FT CARBOHYD 140 140

FT CARBOHYD 143 143

FT CARBOHYD 159 159

FT CARBOHYD 163 163

FT CARBOHYD 188 188

FT CARBOHYD 189 189

FT CARBOHYD 199 199

FT CARBOHYD 209 209

FT CARBOHYD 246 246

FT CARBOHYD 253 253

FT CARBOHYD 274 274

FT CARBOHYD 288 288

FT CARBOHYD 307 307

FT CARBOHYD 307 307

FT CARBOHYD 350 350

FT CARBOHYD 366 366

FT CARBOHYD 372 372

FT CARBOHYD 396 396

FT CARBOHYD 402 402

FT CARBOHYD 408 408

FT CARBOHYD 412 412

FT CARBOHYD 418 418

FT CARBOHYD 423 423

FT CARBOHYD 460 460

FT CARBOHYD 475 475

FT CARBOHYD 622 622

FT CARBOHYD 627 627

FT CARBOHYD 636 636

FT CARBOHYD 648 648

SEQUENCE 867 AA; 98399 MM; 5F231014688B8680 CFC64;

Query Match 87.0%; Score 2723; DA 1; Length 867;

Best Local Similarity 78.6%; Pred. No. 1,2e-212;

Matches 521; Conservative 20; Mismatches 36; Indels 86; Gaps 4;

Qy 2 EKLATVTVYGVVWKEATTLTFCASDAKAYDTEVANNWATHACVTPDPNPQEVLENTE 61

Db 31 EOLWTVTVYGVVWKEATTLTFCASDAKAYDTEVANNWATHACVTPDPNPQEVLENTE 90

Qy 62 HNNMKNNMVBOMODIISLWDSLKPCVKLTPLCGA----- 98

Db 91 KFNMKNNMVBOMEDIISLWDSLKPCVKLTPLCVTLNCIDMGNDTSPNATNTTSSGGE 150

Qy 99 -----GCDT 102

Db 151 KMEKEMKNGSFNITTSIRDKVQKHALFYKHDDVPINNSTRKDNKDNSTRYLISCT 210

Qy 103 SVYTQACRKSIFEPPIHYCAPAGFALLKCDKTPNGGPKCNVSTVOCTHGIRPVSTO 162

Db 211 SVITQACRKSIFEPPIHYCAPAGFALLKCDKCKNGTGCPNVTVOCTHGIRPVSTO 270

Qy 163 LLNLSLAEERVVRSDFTNNAKTIIVOLKESVEINCTRPNNTRKSIIRIQRGPRAPVTT 222

Db 271 LLNLSLAEERVVRSENFNDNAKTIIVOLKESVEINCTRPNNTRKSIIRIQRGPRAPVTT 330

Qy 223 GEIIGDIROACHNISRAKMDTLKQIVIKLREOFENKTIIVFNHSSGGPPEIWMSPNCG 282

Db 331 KOIAGDLQOACHNISRAKMDTLKQIVIKLREOFENKTIIVFNHSSGGPPEIWMSPNCG 390

Qy 283 EFPYCNSTOLFNSWTWN-----NTEGSNNTG-NTTLPCRIKQIIMNOVEGKAMYP 336

Db 391 EFPYCNSTOLFNSWTWN-----NTEGSNNTG-NTTLPCRIKQIIMNOVEGKAMYP 450

Qy 337 IRGOIRGSSNTGLLTRDGGINENGTETFRPGGDMRDNRSELKYKVKVETLEPGLVAP 396

Db 451 IEGQIRGSSNTGLLTRDGGINENGTETFRPGGDMRDNRSELKYKVKVETLEPGLVAP 510

Qy 397 TKCKRVRVOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVOARLLSGIVQOONNLRA 456

Db 511 TKARRVVRREKRAVIGAVELGELGAGSTWG-ASMTLTVAQALLLSGIVQOONNLRA 569  
 Qy 457 IEAQQRMQLTVMGKOLQARVAVERTYLGQOLLIGWGSGKLIICCAVPWNASWSKKS 516  
 Db 570 IEGQOHLQTLTVMGKOLQARVAVERTYLGQOLLIGWGSGKLIICCAVPWNASWSKKS 629  
 Qy 517 LDRIMNNMTWMEWESEIDNYSYITLTLESQNOQEKNOELLDKMASLMMNFDITNW 576  
 Db 630 LEEIMDNMTWMEWESEIDNYSYITLTLESQNOQEKNOELLDKMASLMMNFDITNW 689  
 Qy 577 LMY 579  
 Db 690 LMY 692

RESULT 6  
 ID ENV\_HV1B1 STANDARD; PRT; 856 AA.  
 AC P03375;  
 21-JUL-1986 (Rel. 01, Created)  
 21-JUL-1986 (Rel. 01, Last sequence update)  
 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11678;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8511123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L., Latcheberger J.A., Papas T.S., Chirayil J., Chang N.T., Gallo R.C., Wong-Staal F.;  
 RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RT Nature 313:277-284(1985).  
 RL [2]  
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=90285159; PubMed=2355006;  
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N., Gregory T.J.;  
 RA "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.";  
 RT J. Biol. Chem. 265:10373-10382(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M15654; AAA4205.1; -  
 DR PIR: A03973; VCLTH3.  
 DR HIV: M15654; ENV5BH102.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.  
 KM Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 1 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157

FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 356 356  
 FT CARBOHYD 386 386  
 FT CARBOHYD 392 392  
 FT CARBOHYD 397 397  
 FT CARBOHYD 406 406  
 FT CARBOHYD 448 448  
 FT CARBOHYD 463 463  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 625 625  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 750 750  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97224 MW; 0BF31A18931B27 CRC64;

Query Match 86.9%; Score 2719; DB 1; Length 856;  
 Best Local Similarity 80.0%; Pred. No. 2, 4e-212;  
 Matches 521; Conservative 28; Mismatches 28; Indels 74; Gaps 6;

Qy 2 EKLWVTVYGVVWKEATTLPCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 61  
 Db 32 EKLWVTVYGVVWKEATTLPCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 91  
 Qy 62 HFNWKKNNMEQONEDIIISLMDQSLKCVKLTPLCGA----- 98  
 Db 92 NFNWKKNNMEQONEDIIISLMDQSLKCVKLTPLCVSLKCTDLKNDNTNSSSGRMIMTK 151  
 Qy 99 -----GDPYVITQACPKISFE 115  
 Db 152 GEINCSFNISTSRGVQKEAFYKLDIIPINDTSTYLSNNTSVITQACPKISFE 211  
 Qy 116 PIPHYCAPAGFALIKNDKTPNGKPCQKNSVYQCTHGIRPVVSTOLLNLNGSLAEEBV 175  
 Db 212 PIPHYCAPAGFALIKNDKTPNGKPCQKNSVYQCTHGIRPVVSTOLLNLNGSLAEEBV 271  
 Qy 176 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSHI--GGPAAFTTBEIIGDIOAH 233  
 Db 272 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSHI--GGPAAFTTBEIIGDIOAH 330  
 Qy 234 CNISRAKNDPLKQIVLKEBOF-ENKTIIVFNHSSGDPETVMSFNGGSEFFYCNSVOL 292  
 Db 331 CNISRAKNDPLKQIVLKEBOF-ENKTIIVFNHSSGDPETVMSFNGGSEFFYCNSVOL 390  
 Qy 293 FNSITWN---NTEGSNNTEG-NTITLPCRIKQIINNMQEVGKAYAPPIRGQIRCSSNIT 348  
 Db 391 FNSITWN---NTEGSNNTEG-NTITLPCRIKQIINNMQEVGKAYAPPIRGQIRCSSNIT 450  
 Qy 349 GLILTRDGGINENCTEIFRRGGGMDRDNMSSELYKYVYKIEPIGVAFTCKRRVVRREK 408  
 Db 451 GLILTRDGGINENCTEIFRRGGGMDRDNMSSELYKYVYKIEPIGVAFTCKRRVVRREK 510



GN ENV.  
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
 OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11700;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8511157; PubMed=2982104;  
 RA Mesling M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laasy L.A.,  
 RA Capon D.J.;  
 RT "Nucleic acid structure and expression of the human  
 RT AIDS/lymphadenopathy retrovirus.",  
 RL Nature 313:450-458(1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K02083; AAB59873.1; -.  
 DR EMBL: X01762; CAA25903.1; ALT\_SEQ.  
 DR PIR: A03974; VCLJLV.  
 DR HIV: K02083; ENVSPV22.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KM AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KM Signal.  
 FT SIGNAL. 1 30  
 FT CHAIN 511 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 138 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;  
 Query Match 86.8%; Score 2715; DB 1; Length 856;  
 Best Local Similarity 79.9%; Pred. No. 5e-212;  
 Matches 520; Conservative 27; Mismatches 30; Indels 74; Gaps 6;  
 QY 2 EKLWVTVYGVVPMKEATTTLPFCASDAKAYDTEVHNWATHACVPDPNPOEVLNATE 61  
 DB 32 EKLWVTVYGVVPMKEATTTLPFCASDAKAYDTEVHNWATHACVPDPNPOEVLNATE 91  
 QY 62 HFNWKKNNWYEQMOEDIIISLWDSLRKCVKLTPLCGA----- 98  
 DB 92 NFNWKKNNWYEQMOEDIIISLWDSLRKCVKLTPLCGA----- 98  
 QY 99 -----GCDTSVTOACPKISFE 115  
 DB 152 GEIKNCSFNISTSRGVQKVEAFPYKLDIIPIDNDTSTYLTSCNTSVITQCPKVSFE 211  
 QY 116 PIPHYCAPAGFAILKCNDFPFGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEV 175  
 DB 212 PIPHYCAPAGFAILKCNDFPFGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAEDEV 271  
 QY 176 IRSDFNFTNAKTTIVOLKESVEINCTRPNNTRKSHI--GPGAFYTTSEITGDIRQAH 233  
 DB 272 IRSDFNFTNAKTTIVOLKESVEINCTRPNNTRKSHI--GPGAFYTTSEITGDIRQAH 330  
 QY 234 CNISRAKMNNTLKOIVTKLEQF--ENKTIVENHSSGGDEPIVHNSFGGEFPYCNSTOL 292  
 DB 331 CNISRAKMNNTLKOIVTKLEQF--ENKTIVENHSSGGDEPIVHNSFGGEFPYCNSTOL 390  
 QY 293 FNSFTWNN---NTEGSNNTG--NTITLPCRKIOIINMWQVGMAYAPPIRGQIRCSNIT 348  
 DB 391 FNSFTWNN---NTEGSNNTG--NTITLPCRKIOIINMWQVGMAYAPPIRGQIRCSNIT 450  
 QY 349 GLITTRGCGINEGTEIFRPGCGMDRNMSELYKRVKIEPLGVAFTCKRRVOREK 408  
 DB 451 GLITTRGCGINEGTEIFRPGCGMDRNMSELYKRVKIEPLGVAFTCKRRVOREK 510  
 QY 409 RAVGIGVPLFGPGAAGSTGGAASMTLTVOARLLSGIVQOONNLRAIEAQQMLQLTIV 468  
 DB 511 RAVGIGVPLFGPGAAGSTGGAASMTLTVOARLLSGIVQOONNLRAIEAQQMLQLTIV 570  
 QY 469 WGIKQLQARVLAVERVYGDQOLLGIWCGSGKLTCTAVPNNASMSKSLDRINWNTWME 528  
 DB 571 WGIKQLQARVLAVERVYGDQOLLGIWCGSGKLTCTAVPNNASMSKSLDRINWNTWME 630  
 QY 529 WEREIDNYTSEIYTLIESQOQKQOELELDKMASLWNFDTWLMY 579  
 DB 631 WEREIDNYTSEIYTLIESQOQKQOELELDKMASLWNFDTWLMY 681  
 RESULT 9  
 ENV\_HVLA2 STANDARD; PRT; 855 AA.  
 ID ENV\_HVLA2  
 AC P03378;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [containing: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).  
 OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85090453; PubMed=2578227;  
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,  
 RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,  
 RA Levy J.A., Dina D., Luciw P.A.;  
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus  
 RT (ARV-2).";

RL Science 227:484-492(1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K02007; AAB59882.1; -  
 DR PIR: A03976; VCLA2.  
 DR HIV: K02007; ENV5SF2.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DR Pfam: PF00517; GP41.1.  
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KM Signal.  
 FT CHAIN 1 29  
 FT SIGNAL 30 509  
 FT CHAIN 510 855  
 FT DISULFID 53 73  
 FT DISULFID 118 208  
 FT DISULFID 125 199  
 FT DISULFID 130 155  
 FT DISULFID 221 250  
 FT DISULFID 231 242  
 FT DISULFID 289 333  
 FT DISULFID 380 442  
 FT DISULFID 387 415  
 FT CARBOHYD 87 87  
 FT CARBOHYD 129 129  
 FT CARBOHYD 140 140  
 FT CARBOHYD 154 154  
 FT CARBOHYD 158 158  
 FT CARBOHYD 184 184  
 FT CARBOHYD 190 190  
 FT CARBOHYD 200 200  
 FT CARBOHYD 233 233  
 FT CARBOHYD 244 244  
 FT CARBOHYD 265 265  
 FT CARBOHYD 279 279  
 FT CARBOHYD 292 292  
 FT CARBOHYD 298 298  
 FT CARBOHYD 304 304  
 FT CARBOHYD 334 334  
 FT CARBOHYD 341 341  
 FT CARBOHYD 358 358  
 FT CARBOHYD 364 364  
 FT CARBOHYD 388 388  
 FT CARBOHYD 394 394  
 FT CARBOHYD 400 400  
 FT CARBOHYD 408 408  
 FT CARBOHYD 445 445  
 FT CARBOHYD 458 458  
 FT CARBOHYD 461 461  
 FT CARBOHYD 610 610  
 FT CARBOHYD 615 615  
 FT CARBOHYD 624 624  
 FT CARBOHYD 636 636  
 FT CARBOHYD 815 815  
 FT CARBOHYD 855 AA; 97438 MW; A3BC20573AAC1A2 CR64;  
 SQ SEQUENCE  
 Query Match 86.7%; Score 2714; DB 1; Length 855;  
 Best Local Similarity 79.3%; Pred. No. 6; 1e-212;  
 Matches 518; Conservative 27; Mismatches 30; Indels 78; Gaps 7;

Db 91 NFNWKNMNVEMOEDIIISLMDQSLKPCVKLTPLCGA-----  
 Qy 69 -----GCDTSVITQACP 111  
 Db 151 EIKKCSFNITTSIRDKOKENALFRNLVDVPIDNASTTNTYRLIHCHNSVITQACP 210  
 Qy 112 ISFEPPIPHYCAPAGFALIKNDKTFNGKGPCKAVSTVQCTHGRIPVSTOLLGSLAE 171  
 Db 211 VSFEPIPHYCTPAGFALIKNDKTFNGKGPCKAVSTVQCTHGRIPVSTOLLGSLAE 270  
 Qy 172 EEWVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDIRO 231  
 Db 271 EEWVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDIRK 330  
 Qy 232 AHCNISRAKANDTLKQIYIKRBOF-EKTTIVFNHSSGDDPEIVHNSGCGEFPYCNST 290  
 Db 331 AHCNISRAKANDTLKQIYIKRBOF-EKTTIVFNHSSGDDPEIVHNSGCGEFPYCNST 390  
 Qy 291 QLFNSTWN-NNTGSGNTEGN-TITLPCRIQIINMOEYKAMVAPPIRGIRCSNIT 348  
 Db 391 QLFNSTWN-NNTGSGNTEGN-TITLPCRIQIINMOEYKAMVAPPIRGIRCSNIT 447  
 Qy 349 GLLLTRDGIN-ENGTEIFRPGGDMRDNMSSELYKYVXIEPLGVAPTYCKRRVQRE 407  
 Db 448 GLLLTRDGIN-ENGTEIFRPGGDMRDNMSSELYKYVXIEPLGVAPTYCKRRVQRE 507  
 Qy 408 KRAVCI-GAVLFGPGAAGSTMGASMTLTQARLLSGIVQOONLLRAIEAOQMLQL 466  
 Db 508 KRAVCI-GAVLFGPGAAGSTMGASMTLTQARLLSGIVQOONLLRAIEAOQMLQL 567  
 Qy 467 TWGKIQOAVLAVERYLGPQOLLGICGSGKLICTAVPMNASWSKSLDRIINNTWT 526  
 Db 568 TWGKIQOAVLAVERYLGPQOLLGICGSGKLICTAVPMNASWSKSLDRIINNTWT 627  
 Qy 527 MEWEREIDNYSSEIYTLLEESQNOEKNOELBLDKWASLMNWDITNNWLMY 579  
 Db 628 MEWEREIDNYSSEIYTLLEESQNOEKNOELBLDKWASLMNWDITNNWLMY 680  
 RESULT 10  
 ENV\_HV1Y2 STANDARD; PRT; 843 AA.  
 AC P35961;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retrovirales; Lentivirales.  
 OX NCBI\_TaxID=36377;  
 RN [1]  
 RP MEDLINE=93021387; PubMed=1404605;  
 RX Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,  
 RA Shaw G.M.;  
 RT "Complete nucleotide sequence, genome organization, and biological  
 RT properties of human immunodeficiency virus type 1 in vivo: evidence  
 RT for limited defectiveness and complementation.";  
 RL J. Virol. 66:6587-6600(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M93258; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: H44001; H44001.

```

DR PDB; 1G9N; 27-DEC-00.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00515; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane; Signal;
KM 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEM 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query March 86.5%; Score 2706; DB 1; Length 843;
Best Local Similarity 79.9%; Pred. No. 2.7e-211;
Matches 513; Conservative 25; Mismatches 36; Indels 68; Gaps 5;

```

```

QY 239 AKNDTLKQIVIKLRQF-ENKTIVENHSSGDPPEIWMHSFNGCGEEFYCNSTOLFENSTW 297
DQ 331 TQMENTLEQIAIKLKEKFGNKKTIIRPSSGQPEIETHFNGCGEEFYCNSTOLF--TW 308
QY 298 NNNTGSSNNTGNTTILPCRIKOIINNMQEVGKAMVAPPIRGOIRCSSNTIGLLTRDGG 357
DQ 389 -NDRKLNT-GRNITLPCRIKOIINNMQEVGKAMVAPPIRGOIRCSSNTIGLLTRDGG 446
QY 358 INENGTETFPGGDMDDNRSELVYKVKIPIGVAPTRCKRRVVRKRAVGAVF 417
DQ 447 KDTNTEIFRPGGDMDDNRSELVYKVKIPIGVAPTRCKRRVVRKRAVGAVF 506
QY 418 LGFLGAGSTMGASMTITVQARLLSGIVQOQNNLLRAIEAOQMLQTLTWGIKQLOAR 477
DQ 507 LGFLGAGSTMGASMTITVQARLLSGIVQOQNNLLRAIEAOQMLQTLTWGIKQLOAR 566
QY 478 VLAVERYLDQQLIGWCGSKLICCTAVPMNASMSKSIDRIWNNMTMEBERIDNYT 537
DQ 567 VLAVERYLDQQLIGWCGSKLICCTAVPMNASMSKSIDRIWNNMTMEBERIDNYT 626
QY 538 SEIYTLLESONOERKEDELLDKWASIMNPDITNMLWY 579
DQ 627 HIYSLIEGQNOQEKNEQELLDKWASIMNPDITNMLWY 668

RESULT 11
ENV_HV1H2 STANDARD; PRT; 856 AA.
ID ENV_HV1H2
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; Pubmed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
virus ";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF038399; AAC82596.1; -
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K33; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENVSHXB2.

```

```

DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Glycoprotein; Transmembrane; Signal;
KW 3d-structure.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 156
FT CARBOHYD 160
FT CARBOHYD 186
FT CARBOHYD 197
FT CARBOHYD 230
FT CARBOHYD 234
FT CARBOHYD 241
FT CARBOHYD 262
FT CARBOHYD 276
FT CARBOHYD 289
FT CARBOHYD 295
FT CARBOHYD 301
FT CARBOHYD 332
FT CARBOHYD 339
FT CARBOHYD 356
FT CARBOHYD 366
FT CARBOHYD 392
FT CARBOHYD 397
FT CARBOHYD 406
FT CARBOHYD 448
FT CARBOHYD 463
FT CARBOHYD 611
FT CARBOHYD 616
FT CARBOHYD 624
FT CARBOHYD 637
FT CARBOHYD 674
FT CARBOHYD 750
FT CARBOHYD 816
SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match 86.4%; Score 2702; DB 1; Length 856;
Beet Local Similarity 79.6%; Pred. No. 5,7e-111;
Matches 518; Conservative 28; Mismatches 31; Indels 74; Gaps 6;

2 EKLMTVTVYGVVPMKEATTTLCASDAKAYDTEVHVWATHACVPTDPNPOEVLNENY 61
32 EKLMTVTVYGVVPMKEATTTLCASDAKAYDTEVHVWATHACVPTDPNPOEVLNENY 91
62 HENMKANNVEMOEDDIISLMDQSLKPCVKLTPLCGA----- 98
92 NFNMMKNDVEMOEDDIISLMDQSLKPCVKLTPLCGA----- 151
99 -----GCFTSYITQACPKYSIE 115
152 GEIKNCSFNISTIRGKVQKEAYAFYKLDIIPIDNDTYSYKLTSCNTSYITQACPKVSPE 211
116 PIPHYCAPAGFAILKCNNDKTFNGKPCCKNASTVOCTHGRIPVSTQLLNGLSLAEEDV 175
212 PIPHYCAPAGFAILKCNNDKTFNGKPCCKNASTVOCTHGRIPVSTQLLNGLSLAEEDV 271
176 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTRKSIHI--GPGRAFTTGEIIGDIRQAH 233

```

```

Db 272 IRSVNFDTNAKTIIVQLKESVEINCTRPNNNTRKSIHI--GPGRAFTTGEIIGDIRQAH 330
Qy 234 CNISRAKNDTLCQIVIKLREOF-ENKTIIVFHSSGGDEIYVHSPNCGEFPYCNSTOL 292
Db 331 CNISRAKNDTLCQIVIKLREOF-ENKTIIVFHSSGGDEIYVHSPNCGEFPYCNSTOL 390
Qy 293 FNSTWNN--NTEGSNTEG-NTITLPCRIRKQIIMMOEYGVKAMYAPPIRGIRCSNIT 348
Db 391 FNSTWNN--NTEGSNTEG-NTITLPCRIRKQIIMMOEYGVKAMYAPPIRGIRCSNIT 450
Qy 349 GLLLRDGGINENGRIEIRPGGDMRDWRSLYKTKYKIPPLGVAAPTKCRRVYQREK 408
Db 451 GLLLRDGGINENGRIEIRPGGDMRDWRSLYKTKYKIPPLGVAAPTKCRRVYQREK 510
Qy 409 RAVGIGAVELGFLGAGSTMGAASTMTLYVQARLLSGIYVQOQNNLIRALEAOQHLQLTV 468
Db 511 RAVGIGAVELGFLGAGSTMGAASTMTLYVQARLLSGIYVQOQNNLIRALEAOQHLQLTV 570
Qy 469 WGIKQLQARVLAVERYLDQQLIGWCSGKLICTTAVPMNWSNKSIDRIYNNNTWME 528
Db 571 WGIKQLQARVLAVERYLDQQLIGWCSGKLICTTAVPMNWSNKSIDRIYNNNTWME 630
Qy 529 WEREIDNTTSBYTLIESQNOQERNEDELELDYKASIMNFDITNMLWY 579
Db 631 WEREIDNTTSBYTLIESQNOQERNEDELELDYKASIMNFDITNMLWY 681

RESULT 12
ENV_HV1H3 STANDARD; PRT; 856 AA.
AC P04624;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
OX NCBI_Taxid=11707;
RN [1]
RP MEDLINE=85228248; PubMed=2988795;
RA Crowl R., Ganguily K., Gordon M., Conroy R., Schaber M., Kramer R.,
Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HIV-1 env gene products synthesized in E. coli are recognized by
antibodies present in the sera of AIDS patients.";
RL Cell 41:979-986 (1985).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M14100; AAA44679.1; -
CC PDB; 1DAU; 17-OCT-01.
CC PDB; 1DAV; 17-OCT-01.
CC HIV; M14100; ENVSHXB3.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW Aids; Coat protein; Glycoprotein; Transmembrane; Signal;
KW 3d-structure.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 156
FT CARBOHYD 160
FT CARBOHYD 186
FT CARBOHYD 197
FT CARBOHYD 230
FT CARBOHYD 234
FT CARBOHYD 241
FT CARBOHYD 262
FT CARBOHYD 276
FT CARBOHYD 289
FT CARBOHYD 295
FT CARBOHYD 301
FT CARBOHYD 332
FT CARBOHYD 339
FT CARBOHYD 356
FT CARBOHYD 366
FT CARBOHYD 392
FT CARBOHYD 397
FT CARBOHYD 406
FT CARBOHYD 448
FT CARBOHYD 463
FT CARBOHYD 611
FT CARBOHYD 616
FT CARBOHYD 624
FT CARBOHYD 637
FT CARBOHYD 674
FT CARBOHYD 750
FT CARBOHYD 816
SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

```



```

FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MM; 3373C8BB84C1AFC CRC64;

```

```

Query Match 86.3%; Score 2699; DB 1; Length 856;
Best Local Similarity 79.6%; Pred. No. 1e-210;
Matches 518; Conservative 26; Mismatches 33; Indels 74; Gaps 6;

```

```

QY 2 EKLWTVYGVVWKEATTTLLFCASADAYDDEVNVAHACVPTDPDPEVLENTTE 61
DB 32 EKLWTVYGVVWKEATTTLLFCASADAYDDEVNVAHACVPTDPDPEVLENTTE 91
QY 62 HFNWKNMNVROMEDIIISLMDQSLKPCVKLTPLCGA----- 98
QY 92 NFNWKNMNVROMEDIIISLMDQSLKPCVKLTPLCGA----- 151
QY 99 -----GGDTVITQACPKISFE 115
DB 152 GEIKKCSNISTSRGKVOKEYAFYKLDIIPIDNDTTSYLTSCNTSVITQACPKISFE 211
QY 116 PIPHYCAPAGFAILKCNKDTFNGKPKKNVSTVOCTGIRVAVSTOLLNGLAEEBIV 175
DB 212 PIPHYCAPAGFAILKCNKDTFNGKPKKNVSTVOCTGIRVAVSTOLLNGLAEEBIV 271
QY 176 IRSDFNNAKTIIVQLKESVEINCTRPNNNTKRSIHI--GGRAFYTTGELIGDIRQAH 233
DB 272 IRSVFNNAKTIIVQLNTSVEINCTRPNNNTKRSIHI--GGRAFYTTGELIGDIRQAH 330
QY 234 CNISAKKNDTLKQIVIKLRQF--ENKTIIVFNHSGGDEIYHMFNFGGGEFFYCNSTQL 292
DB 331 CNISAKKNDTLKQIVIKLRQF--ENKTIIVFNHSGGDEIYHMFNFGGGEFFYCNSTQL 390
QY 293 FNSTWNN--NTEGSNNTG--NTITLPCRIKQIIMMOBVGKAMAPYIRGQIRCSNIT 348
DB 391 FNSTWNN--NTEGSNNTG--NTITLPCRIKQIIMMOBVGKAMAPYIRGQIRCSNIT 450
QY 349 GLLTLTRDGGINENGTEIFRPGGDMRDMNRSELVYKVKVILPGLVAPLTKCRVVOREK 408
DB 451 GLLTLTRDGGINENGTEIFRPGGDMRDMNRSELVYKVKVILPGLVAPLTKCRVVOREK 510

```

```

QY 409 RAVGIGAVFLFGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLTV 468
DB 511 RAVGIGAVFLFGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLTV 570
QY 469 WGIKQLQARVAVRYLYDQQLIGWCCSGKLCCTTAVPNVNASWNSKSLDRINNTWME 528
DB 571 WGIKQLQARVAVRYLYDQQLIGWCCSGKLCCTTAVPNVNASWNSKSLDRINNTWME 630
QY 529 MERIDNVTSEIYTLIEESONQOQKNEOFTELDKASLNNWPTTWLWY 579
DB 631 WDREINNTSLHSLIEESONQOQKNEOFTELDKASLNNWPTTWLWY 601

RESULT 13
ENV_HV1B8 STANDARD; PRT; 851 AA.
ID ENV_HV1B8
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP120 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Joseph S.F., Doran E.R., Ratafski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Pappas T.S., Graybe J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284 (1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; K02011; AAA44661.1; -.
DR PDB; 1A1K; 16-JUN-97.
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QO3; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENV5B8.
DR GlycoSuiteDB; P04582; -.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 506
FT CHAIN 507 851
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 440
FT DISULFID 385 413
FT CARBOHYD 88 88
FT DISULFID 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).

```







FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 96938 MW; 0C241332CF7E6687 CRC64;

## Query Match

86.1%; Score 2693; DB 1; Length 856;

Best Local Similarity 79.1%; Pred. No. 3.1e-210;  
Matches 515; Conservative 29; Mismatches 33; Indels 74; Gaps 6;

QY 2 EKLMTVTVYGVVPMWEATITTLFCASDAKAYDTEVNNVWATHACVPTDPNPOEVLENTE 61  
DB 32 EKLMTVTVYGVVPMWEATITTLFCASDAKAYDTEVNNVWATHACVPTDPNPOEVLENTE 91  
QY 62 HFNMMKNVNEQMOEDIIISLMDPSLKPCYKLPPLGCA----- 98  
DB 92 HFNMMKNVNEQMOEDIIISLMDPSLKPCYKLPPLGCA----- 151  
QY 99 -----GCDTSVITQACPKISPE 115  
DB 152 GEIKNCSFNISTIRGKVOKEVAFYKHDIIPIDNDTSVTLTSCNTSVITQACPKISPE 211  
QY 116 PIPHYCAPAGFALLKCNKTFNGKPCQNVSTVOCTHGIRPVSTQLNLNGSLAEVY 175  
DB 212 PIPHYCAPAGFALLKCNKTFNGKPCQNVSTVOCTHGIRPVSTQLNLNGSLAEVY 271  
QY 176 IRSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSIH--GPRAFYTTGELIIGDIRAH 233  
DB 272 IRSAITLDVKTIIIVOLNOSVEINCTRPNNNTRKRIORGPRTFTYIGK--IGNRORAH 330  
QY 234 CNISRAKNNDTLKOIIVIKLREOF-ENKTIIVFNHSSGDEPEIVHSHFCGEEFFYCNSIOL 292  
DB 331 CNISRAKNNDTLKOIIVIKLREOF-ENKTIIVFNHSSGDEPEIVHSHFCGEEFFYCNSIOL 390  
QY 293 FNSTNN--NTEGSSNTEG-NTTILPCRIKOIINNMOEVGKAMVAPPIRGOIRCSSNIT 348  
DB 391 FNSTNN--NTEGSSNTEG-NTTILPCRIKOIINNMOEVGKAMVAPPIRGOIRCSSNIT 450  
QY 349 GLLTRDGINENGTEIFRPGGDMRDNRSLEYKVKVIEPIGVAPTCKRVRVQREK 408  
DB 451 GLLTRDGINENGTEIFRPGGDMRDNRSLEYKVKVIEPIGVAPTCKRVRVQREK 510  
QY 409 RAVGIGAVELGAGASTGASMTLTVQARLLLSGIVOQONNLRAIEAOQMLQLTV 468  
DB 511 RAVGIGAVELGAGASTGASMTLTVQARLLLSGIVOQONNLRAIEAOQMLQLTV 570  
QY 468 WGIKQLOARVLAVERLYGDOQLIGWCGGKLICTTAVPMAASWSNKSIDRIWNNMTME 528  
DB 571 WGIKQLOARVLAVERLYGDOQLIGWCGGKLICTTAVPMAASWSNKSIDRIWNNMTME 630  
QY 529 WREIDNTYSEIYTLIEESONOEKNEBELLDKASLMMNPDTNMLMY 579  
DB 631 WREIDNTYSEIYTLIEESONOEKNEBELLDKASLMMNPDTNMLMY 681

Search completed: December 12, 2003, 12:30:48  
Job time : 11.7179 secs

**THIS PAGE BLANK (US:70)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:24:08 ; Search time 37.2225 Seconds

(without alignments)  
2665.163 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351  
Sequence: 1 VEKLMVTVYGVVPWKREATT.....EIDKWSASIMNFDITKMLMY 625

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
A.GeneSeq\_19Jun03:\*

1:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	99.8	625	22 AAB61507	HIV-1 delTAV3 SOS
2	3295	98.3	643	22 AAB61505	HIV-1 SOS gp140 g1
3	2968.5	88.6	850	16 AAR67724	gp120 from the HIV
4	2965	88.5	883	22 AAB82761	Ancestral HIV-1 gr
5	2938.5	87.7	855	18 AAM1581	Human Immunodefici
6	2938.5	87.7	855	20 AAM88113	Env protein of the
7	2928	87.4	847	21 AAY97073	Variant HIV-1 SF16
8	2923	87.2	579	22 AAB61506	HIV-1 delTAV12* S
9	2923	87.2	842	24 AAB61505	Human immunodefici

10	2918	87.1	842	23 AAB606211	HIV Env isolate SF
11	2898.5	86.5	868	27 AAB60422	Sequence of LAV v1
12	2895.5	86.4	868	23 AAO13389	Lymphadenopathy-as
13	2894.5	86.4	868	7 AAB60063	HIV virus env gene
14	2887	86.2	856	22 AAB85999	Amino acid sequenc
15	2887	86.2	863	14 AAR43869	HTLV-III ENV-LOR g
16	2886	86.1	854	21 AAB10697	HIV-1 env protein.
17	2886	86.1	854	21 AAB10053	HIV-1 coat protein
18	2886	86.1	854	22 AAB86199	HIV gp41 DNA SEQ I
19	2886	86.1	854	22 AAB67277	Protein encoded by
20	2886	86.1	854	23 ABR73663	HIV-1 NL4-3 gp120
21	2885	86.1	856	7 AAB60131	Sequence of the A1
22	2884.5	86.1	851	7 AAB80967	HIV protein HT6.
23	2883	86.0	856	7 AAB61514	Sequence of envelo
24	2883	86.0	856	20 AAM89325	HIV-1 env protein
25	2881.5	86.0	901	8 AAB70665	Sequence encoded b
26	2879	85.9	856	13 AAR25940	Modified HIV env g
27	2878	85.9	856	21 AAY97072	Wild type HIV-1 HX
28	2877	85.9	863	7 AAB60349	HTLV-III virus (HI
29	2876	85.8	857	16 AAB67725	gp120 from the HIV
30	2871	85.7	863	14 AAR43873	HTLV-III ENV-LOR g
31	2871	85.7	1188	17 AAR94622	HIV-1 gp160-NY5 p5
32	2870	85.6	856	18 AAW11579	Human Immunodefici
33	2870	85.6	856	20 AAW88111	Predicted amino ac
34	2869.5	85.6	751	10 AAB90604	Human immunodefici
35	2867	85.6	856	22 AAB45697	HIV envelope prote
36	2861	85.4	857	12 AAR14904	HIV-1/IIIB env clo
37	2861	85.4	857	18 AAW11580	Human Immunodefici
38	2861	85.4	857	20 AAW88112	Predicted amino ac
39	2858.5	85.3	855	19 AAW43069	HIV-1 gp120 protei
40	2854.5	85.2	880	18 AAW23333	Human immunodefici
41	2854.5	85.2	880	20 AAW73332	HIV envelope prote
42	2854.5	85.2	880	21 AAB14844	HIV-1 gp120 protei
43	2854	85.2	858	19 AAW43067	HIV-1 gp120 protei
44	2853	85.1	856	20 AAW90178	HTLV-III env prote
45	2851.5	85.1	855	12 AAR14905	HIV-1 BA-L clone-e

## ALIGNMENTS

RESULT 1  
AAB61507  
ID AAB61507 standard; Protein; 625 AA.  
AC AAB61507;  
XX  
DT 05-APR-2001 (first entry)  
XX  
DE HIV-1 delTAV3 SOS gp140 glycoprotein.  
XX  
KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;  
KW gp140.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200100648-A1.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-US17267.  
XX  
PR 25-JUN-1999; 99US-0340992.  
XX  
XX (PROG-) PROGENICS PHARM INC.  
PA (AARO-) AARON DIMOND AIDS RES CENT.  
XX  
PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;  
XX WPI; 2001-122993/13.  
XX N-PSDB; AAF28583.  
XX  
PT New viral envelope proteins, useful for producing vaccines to treat

PT human immunodeficiency virus-1 infections, comprises amino acid  
 PT sequence mutations such that viral transmembrane-surface protein  
 PT complex is more stable  
 XX  
 XX  
 PS Disclosure; Fig 15; 109pp; English.

CC The present invention relates to a viral envelope protein. The viral  
 CC envelope protein comprises a viral surface protein (e.g. glycoprotein  
 CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in  
 CC which the viral envelope protein contains one or more amino acid sequence  
 CC mutations that enhance the stability of the complex formed between the  
 CC viral surface and transmembrane proteins. The viral envelope protein can  
 CC be used in the treatment of viral infection e.g. HIV-1 infection. The  
 CC present sequence is HIV-1 deltaV3 SOS gp140, which was used in the  
 CC present invention.

CC Sequence 625 AA;

Query Match 99.8%; Score 3345; DB 22; Length 625;

at Local Similarity 100.0%; Pred. No. 2.5e-188; Mismatches 0; Indels 0; Gaps 0;

atches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEKLMVTVVYGVVPMWKEATTTLFCASDAKAYDTEVHNWMAHACVPTDPPNPOEVLENT 60

DB 1 VEKLMVTVVYGVVPMWKEATTTLFCASDAKAYDTEVHNWMAHACVPTDPPNPOEVLENT 60

QY 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 120

DB 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 120

QY 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180

DB 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180

QY 181 IPIHYCAPAGFALLKNDKTNPKXPCKNVSTXCTGIRPVVSTOLLNGLAEEVVI 240

DB 181 IPIHYCAPAGFALLKNDKTNPKXPCKNVSTXCTGIRPVVSTOLLNGLAEEVVI 240

QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRAHCHISRAKMDTLKQIVIKLRE 300

DB 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRAHCHISRAKMDTLKQIVIKLRE 300

QY 301 QFENKTIIVFNHSSGGDEPEIVMHSFNGGEPFYCNSTOLFNTNNTGSGNNTGNTITL 360

DB 301 QFENKTIIVFNHSSGGDEPEIVMHSFNGGEPFYCNSTOLFNTNNTGSGNNTGNTITL 360

QY 361 PCRIKOIINMVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 420

DB 361 PCRIKOIINMVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 420

QY 421 DNMRSELYKYVAVIPELGVAPTRCKRRVVOREKRAVGIGAVPLGFGAAGSTMGAASMT 480

DB 421 DNMRSELYKYVAVIPELGVAPTRCKRRVVOREKRAVGIGAVPLGFGAAGSTMGAASMT 480

QY 481 LTVQARLLSGIIVQOQNNLLRAIEAQQRMLQLTWGIKQIOLARVLAVERVYGDQDLGIM 540

DB 481 LTVQARLLSGIIVQOQNNLLRAIEAQQRMLQLTWGIKQIOLARVLAVERVYGDQDLGIM 540

QY 541 GCSGKLICTTAVPNNASWNSKSLDRINNMNTMWEREIDVYSEIYTLIEESQNOQCKN 600

DB 541 GCSGKLICTTAVPNNASWNSKSLDRINNMNTMWEREIDVYSEIYTLIEESQNOQCKN 600

QY 601 EOELELDKASLNMWPDITKMWY 625

DB 601 EOELELDKASLNMWPDITKMWY 625

RESULT 2

AAB61505 ID AAB61505 standard; Protein; 643 AA.

XX AC AAB61505;

XX

DT 05-APR-2001 (first entry)

XX HIV-1 SOS gp140 glycoprotein.

XX gp120; anti-HIV, gene therapy; vaccine; viral envelope protein; gp41;

XX gp140.

XX Human immunodeficiency virus type 1.

XX WO2001/00648-A1.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17267.

XX 25-JUN-1999; 99US-0340992.

XX (PROG-) PROGENICS PHARM INC.

XX (AARO-) AARON DIAMOND AIDS RES CENT.

XX Binley JM, Schuelke N, Olson WC, Madden PJ, Moore JP;

XX WPI; 2001-122993/13.

XX N-PSDB; AAF28581.

XX New viral envelope proteins, useful for producing vaccines to treat

XX human immunodeficiency virus-1 infections, comprises amino acid

XX sequence mutations such that viral transmembrane-surface protein

XX complex is more stable

XX Disclosure; Fig 13; 109pp; English.

XX The present invention relates to a viral envelope protein. The viral

XX envelope protein comprises a viral surface protein (e.g. glycoprotein

XX gp120) and a corresponding viral transmembrane protein (e.g. gp41), in

XX which the viral envelope protein contains one or more amino acid sequence

XX mutations that enhance the stability of the complex formed between the

XX viral surface and transmembrane proteins. The viral envelope protein can

XX be used in the treatment of viral infection e.g. HIV-1 infection. The

XX present sequence is HIV-1 SOS gp140, which was used in the present

XX invention.

XX Sequence 643 AA;

Query Match 98.3%; Score 3295; DB 22; Length 643;

at Local Similarity 96.0%; Pred. No. 2.2e-185; Mismatches 8; Indels 18; Gaps 1;

Matches 617; Conservative 0; Mismatches 8; Indels 18; Gaps 1;

QY 1 VEKLMVTVVYGVVPMWKEATTTLFCASDAKAYDTEVHNWMAHACVPTDPPNPOEVLENT 60

DB 1 VEKLMVTVVYGVVPMWKEATTTLFCASDAKAYDTEVHNWMAHACVPTDPPNPOEVLENT 60

QY 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 120

DB 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 120

QY 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180

DB 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180

QY 181 IPIHYCAPAGFALLKNDKTNPKXPCKNVSTXCTGIRPVVSTOLLNGLAEEVVI 240

DB 181 IPIHYCAPAGFALLKNDKTNPKXPCKNVSTXCTGIRPVVSTOLLNGLAEEVVI 240

QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRAHCHISRAKMDTLKQIVIKLRE 300

DB 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRAHCHISRAKMDTLKQIVIKLRE 300

QY 283 SRKAMNDTLKQIVIKLEQFENKTIIVFNHSSGGDEPEIVMHSFNGGEPFYCNSTOLFNT 342

DB 301 SRKAMNDTLKQIVIKLEQFENKTIIVFNHSSGGDEPEIVMHSFNGGEPFYCNSTOLFNT 360

QY 343 WNNNTGSGNNTGNTITLPCRIKOIINMVEQWEDIIISLMDOSLKPCKVLTPLCYTLNCKDVNATNTNDSGTMERG 402

```

Db      |||||
361  MNWNNEGSNNTGNTITTLPCR|KOIIMMOEVBKAMYPPIRGOIRCSSNITGLLTRG 420
Qy      403  GINENGTEIFRPGGDMRDNRMSSELYKYVVKIEPLGVAPTKCKRVVOREKRAVIGAV 462
Db      421  GINENGTEIFRPGGDMRDNRMSSELYKYVVKIEPLGVAPTKCKRVVOREKRAVIGAV 480
Qy      463  FLGFLGAAGSTWGAASMTLTVQARLLLSGIVOOQNNLLRAIEAQRMLQLTWGIKQLOA 522
Db      481  FLGFLGAAGSTWGAASMTLTVQARLLLSGIVOOQNNLLRAIEAQRMLQLTWGIKQLOA 540
Qy      523  RVLAVERYIGDQQLGIGWCGSGKLTCTTAVPNNASWSNKSJLDRIINNMTMEMEREIDNY 582
Db      541  RVLAVERYIGDQQLGIGWCGSGKLTCTTAVPNNASWSNKSJLDRIINNMTMEMEREIDNY 600
Qy      583  TSEIYTLIEESONOEKNEQELLEDKMASLNNPDTIKLWY 625
Db      601  TSEIYTLIEESONOEKNEQELLEDKMASLNNPDTIKLWY 643

```

LT 3  
57724

ID AAR67724 standard; Protein; 850 AA.

AC AAR67724;

DT 25-MAR-2003 (updated)  
DT 07-SEP-1995 (first entry)

DE gp120 from the HIV GNE8 isolate.

KM HIV; human immunodeficiency virus; gp120; glycoprotein;  
KW GNE clone; GNE8 isolate.

OS Human immunodeficiency virus type 1.

PN WO9428929-A1.

PD 22-DEC-1994.

PF 07-JUN-1994; 94WO-US06036.

PR 07-JUN-1993; 93US-0072833.

PA (GETH ) GENENTECH INC.

PI Berman PW, Nakamura GR;

XX MPI; 1995-036112/05.  
XX N-PEDB; AAQ76018.

PT Use of HIV gp 120 polypeptide(s) - for developing probes for the  
PT analysis, prevention and therapy of HIV infection

PS Claim 23; Page 27-30; 108pp; English.

CC This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate  
CC of HIV-1 (human immunodeficiency virus type 1). There are neutralising  
CC epitopes in the V2 and C4 domains of gp120, in addition to the  
CC neutralising epitopes in the V3 domain. Although the amino acid sequences  
CC of the neutralising epitopes in the V2, V4 and C4 domains are variable,  
CC the amount of variation is highly constrained. This facilitates the  
CC design of HIV subunit vaccines that can induce antibodies that neutralise  
CC the most common HIV strains for a given geographic region. This invention  
CC provides a multivalent gp120 subunit vaccine where the gp120 present in  
CC the vaccine is from at least two HIV isolates which have different amino  
CC acid sequences for a neutralising epitope in these regions.  
CC (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 850 AA;

Query Match 88.6%; Score 2968.5; DB 16; Length 850;  
Best Local Similarity 86.9%; Pred. No. 3.8e-166;

Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

```

Qy      2  EKLAVTVYGVVPMKEATTLTFCASDAXAYDEVHNVNATHACVPTDPPOVLENTYE 61
Db      31  EKLAVTVYGVVPMKEATTLTFCASDAXAYDEVHNVNATHACVPTDPPOVLENTYE 90
Qy      62  HENMMKNMVEQMOBIIISLMDOSLPCVKLTPLCVTLNCKDV-NATNTNDSBETMERG 120
Db      91  NFNMMKNMVEQMHEDIIISLMDOSLPCVKLTPLCVTLNCTDLKNATNTSSWGMERG 150
Qy      121  EIKNSFNITTSIRDEQVEYALFYKLDVYXIDNNNTSYRLISCTSVYTOACPKISFEP 180
Db      151  EIKNSFNITTSIRDKMKREYALFYKLDVVPIDNNNTSYRLISCVTSVYTOACPKYSFEP 210
Qy      181  IPIHFCAPAGFAILKCNKTFNGKKPCKNVSTXOCTHGIRPVVSTOLLNGLABEYVI 240
Db      211  IPIHFCAPAGFAILKCRDKKFNCTPCTNVSTVQCTHGIRPVVSTOLLNGLABEYVI 270
Qy      241  RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHQNI 282
Db      271  RSANFSDNAKTIIVOLNESVEINCTRPNNNTRRSIHIGRAFYATGEIIGDIRQAHQNL 330
Qy      283  SPAKNDTLKQIVTLREQFENKTIIVFNHSSGDEIYVHNSGCGEPFYCNSTOLFNST 342
Db      331  SSTKMNNTLKQIVTLREHP-NKTIIVFNHSSGDEIYVHNSGCGEPFYCNSTOLFNST 389
Qy      343  WN-----NNTGSSNTEGNTITLPCR|KOIIMMOEVBKAMYPPIRGOIRCSSNITGL 397
Db      390  WNTYTTWNTNBSNNT-GRNITLQCR|KOIIMMOEVBKAMYPPIRGOIRCSSNITGL 448
Qy      398  LTRDGINENGTEIFRPGGDMRDNRMSSELYKYVVKIEPLGVAPTKCKRVVOREKRAV 457
Db      449  LTRDGG-NNSETEIFRPGGDMRDNRMSSELYKYVVKIEPLGVAPTKCKRVVOREKRAV 507
Qy      458  GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVOOQNNLLRAIEAQRMLQLTWGI 517
Db      508  GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVOOQNNLLRAIEAQRMLQLTWGI 567
Qy      518  KQIARVLAVERYLDDQQLGIGWCGSGKLTCTTAVPNNASWSNKSJLDRIINNMTMEMER 577
Db      568  KQIARVLAVERYLDDQQLGIGWCGSGKLTCTTAVPNNASWSNKSJLDRIINNMTMEMER 627
Qy      578  EIDNTSEIYTLIEESONOEKNEQELLEDKMASLNNPDTIKLWY 625
Db      628  EIDNTSEIYTLIEESONOEKNEQELLEDKMASLNNPDTIKLWY 675

```

RESULT 4

AAB82761  
ID AAB82761 standard; Protein; 883 AA.

AC AAB82761;

DT 29-OCT-2001 (first entry)

DE Ancestral HIV-1 group M, subtype B gp160 protein.

KM HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.

OS Human immunodeficiency virus type 1.

PN WO200160838-A2.

PD 23-AUG-2001.

PF 16-FEB-2001; 2001WO-US05288.

PR 18-FEB-2000; 2000US-0183659.

PA (UNITW ) UNIV WASHINGTON.

PI Mullins JT, Rodrigo AG, Learn GH, Li F;





Query Match 87.7%; Score 2938.5; DB 18; Length 855;  
 Best Local Similarity 85.1%; Pred. No. 2.2e-164;  
 Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

```

QY 2 EKLMTVYVGVVWMEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENTYE 61
DB 31 EKLMTVYVGVVWMEATTTTFCASDRKAYDTEVHNWATHACVPTDPNPQEVLEKNVTE 90
QY 62 HFNMMKNNWVEMQMODIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSRG 115
DB 91 NFNMMKNNWVEMQMODIISLMDQSLKPCVKLTPLCVTLNCKDVLRNATNGNDNTNTSSSG 150
QY 116 TMERGEIKNCSFNITTSIRDEVQKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
DB 151 MWGGEEMKNCSEFNITNTIRGKVEYALFYKLDIAPIDNNSNNRYRLISCNSTSVITQACP 210
QY 175 KISFEPIDPHYCAPAGFALCKNDKTFNGKXPCXNVSTXOCTHGIRPVVSTOLLNGSLA 234
DB 211 KVSFEPIDPHYCAPAGFALCKDKKFKNGKGTNVSTVQCTHGIRPVVSTOLLNGSLA 270
QY 235 BEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNG-----AGDIR 276
DB 271 BEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNTKRSIHIGGRAPYTTGELIGDIR 330
QY 277 OAHCHNISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDPRIVHSHFNGGEPFYCNST 336
DB 331 OAHCHNISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDPRIVHSHFNGGEPFYCNST 390
QY 337 QLFNSTNNNTGSGNNTGNTITLPCRIKQIIMMOEYKAMYPAPIRGOIRCSSNITGL 396
DB 391 QLFNSTNNNTGSGNNTGNTITLPCRIKQIIMMOEYKAMYPAPIRGOIRCSSNITGL 450
QY 397 LITRDGGINENGTEIFRPGGDMRDNMSELYKYKVVKEIPLGVAPTCRKRRVOREKXA 456
DB 451 LITRDGGEEDNKTIEFRPGGDMRDNMSELYKYKVVKEIPLGVAPTCRKRRVOREKXA 510
QY 457 VGIGAVPLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQOHLDTLWVG 516
DB 511 VGIGAVPLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQOHLDTLWVG 570
QY 517 IKOLQARVLAVERYLGDQDLGIGWCSGLICTTAVPNNASWSNKSLSLRINNMNTWMEWE 576
DB 571 IKOLQARVLAVERYLGDQDLGIGWCSGLICTTAVPNNASWSNKSLSLRINNMNTWMEWE 630
QY 577 REIDNVTSEIYTLIESONQOEKNEOELELDKMASLMMWPDITKMLWY 625
DB 631 REIDNVTSEIYTLIESONQOEKNEOELELDKMASLMMWPDITKMLWY 679

```

JLT 6  
 ID AAW88113 standard; Protein; 855 AA.  
 AAW88113;  
 AC AAW88113;  
 DT 09-APR-1999 (first entry)  
 DE Env protein of the BA-L strain of Human immunodeficiency virus type 1.  
 XX HIV-1; HIV-1 strain BA-L; env protein; vaccine;  
 KM immunotherapy; HIV infection; immunogen; HIV-1 diagnosis.  
 OS Human immunodeficiency virus type 1.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 11 /note= "Gln encoded by CG"  
 XX  
 PN US869313-A.  
 PD 09-FEB-1999.  
 XX  
 PF 14-MAY-1996; 96US-0647714.

XX 17-OCT-1990; 90US-0599491.  
 PR 25-FEB-1993; 93US-0022835.  
 PR 15-FEB-1995; 95US-0388809.  
 PR 14-MAY-1996; 96US-0647714.  
 XX  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PI Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD,  
 PI Popovic M, Reitz MS;  
 XX WPI: 1999-152779/13.  
 DR N-PSDB; AAX04767.  
 PT DNA encoding env protein of the human immune deficiency virus  
 PT isolate BA-L - useful for producing protein for use in vaccines, as  
 PT assay reagent and to generate antibodies  
 XX  
 PS Example 1; Fig 9A-C; 87pp; English.  
 CC (ATCC 40890) strain of Human immunodeficiency virus type 1 (HIV-1)  
 CC strain MN-STL. BA-L is more typical of United States isolates of  
 CC HIV-1 than previously known strains. Recombinant, complete env protein  
 CC of the BA-L strain is used as a vaccine component and for immunotherapy  
 CC of existing HIV infections, to detect HIV-specific antibodies, e.g. in  
 CC donated blood, and as an immunogen to raise specific antibodies, for  
 CC HIV-1 diagnosis.

SQ Sequence 855 AA;  
 Query Match 87.7%; Score 2938.5; DB 20; Length 855;  
 Best Local Similarity 85.1%; Pred. No. 2.2e-164;  
 Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

```

QY 2 EKLMTVYVGVVWMEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENTYE 61
DB 31 EKLMTVYVGVVWMEATTTTFCASDRKAYDTEVHNWATHACVPTDPNPQEVLEKNVTE 90
QY 62 HFNMMKNNWVEMQMODIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSRG 115
DB 91 NFNMMKNNWVEMQMODIISLMDQSLKPCVKLTPLCVTLNCKDVLRNATNGNDNTNTSSSG 150
QY 116 TMERGEIKNCSFNITTSIRDEVQKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
DB 151 MWGGEEMKNCSEFNITNTIRGKVEYALFYKLDIAPIDNNSNNRYRLISCNSTSVITQACP 210
QY 175 KISFEPIDPHYCAPAGFALCKNDKTFNGKXPCXNVSTXOCTHGIRPVVSTOLLNGSLA 234
DB 211 KVSFEPIDPHYCAPAGFALCKDKKFKNGKGTNVSTVQCTHGIRPVVSTOLLNGSLA 270
QY 235 BEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNG-----AGDIR 276
DB 271 BEEVIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNTKRSIHIGGRAPYTTGELIGDIR 330
QY 277 OAHCHNISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDPRIVHSHFNGGEPFYCNST 336
DB 331 OAHCHNISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDPRIVHSHFNGGEPFYCNST 390
QY 397 LITRDGGINENGTEIFRPGGDMRDNMSELYKYKVVKEIPLGVAPTCRKRRVOREKXA 456
DB 451 LITRDGGEEDNKTIEFRPGGDMRDNMSELYKYKVVKEIPLGVAPTCRKRRVOREKXA 510
QY 457 VGIGAVPLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQOHLDTLWVG 516
DB 511 VGIGAVPLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQOHLDTLWVG 570
QY 517 IKOLQARVLAVERYLGDQDLGIGWCSGLICTTAVPNNASWSNKSLSLRINNMNTWMEWE 576

```

Db 571 IKQLQARLVAERYLRDQQLIGWCGSKGLICTTAVPNNASMSKSLNKIKWNTMTIEWMD 630

Qy 577 REINNTSEIYTLIEESONQOEKNEQELLELDKNASLNNWFDITKMLWY 625

Db 631 REINNTSIYSLIEESONQOEKNEQELLELDKNASLNNWFDITKMLWY 679

RESULT 7

AA97073

AA97073 standard; protein; 847 AA.

AC AA97073;

DT 31-OCT-2000 (first entry)

DE Variant HIV-1 SF162 Env gp160.

KM HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env; CD4 binding region; VI/V2 loop; bridging sheet.

Human immunodeficiency virus type 1 isolate SF162.

MO200039303-A2.

06-JUL-2000.

30-DEC-1999; 99WO-US31272.

31-DEC-1998; 98US-0114495.

29-SEP-1999; 99US-0156670.

(CHIR) CHIRON CORP.

Barnett S, Hartog K, Martin E;

WPI; 2000-465745/40.

Novel modified HIV Env polypeptides useful as immunizing agents and for preparing a vaccine to elicit an immune response against a broad range of HIV subtypes

Claim 5; Page 115-117; 139pp; English.

Novel immunogenic modified human immunodeficiency virus (HIV) envelope (Env) polypeptides having an amino acid deleted or replaced in the region corresponding to residues 420-436 or 119-123 and 199-210 relative to isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV strain SF162, with numbering relative to isolate HXB-2. The Env polypeptides are modified so as to expose at least part of the CD4 binding region. The modified HIV Env polypeptides, coding polynucleotides and constructs, further comprising an adjuvant, are used for inducing an immune response in an individual. The method involves administering a first composition comprising a polynucleotide encoding the Env polypeptide in a priming step and administering a second composition comprising a modified Env polypeptide as a booster in an amount sufficient to induce an immune response in the individual. The first and/or second composition further comprises an adjuvant (claimed). The intracellularly produced Env polypeptides can be used for a number of diagnostic and therapeutic purposes to determine the presence of reactive antibodies and/or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

Sequence 847 AA;

Query Match 87.4%; Score 2928; DB 21; Length 847;

Best Local Similarity 86.0%; Pred. No. 9e-164;

Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

Qy 1 VEKLMVTVYGVVPMKEATITLFCASDAKAYDTEVHNWMTACVPTPNQEVLENTV 60

Db 30 VEKLMVTVYGVVPMKEATITLFCASDAKAYDTEVHNWMTACVPTPNQEVLENTV 89

Qy 61 EHFNMNNNNVQOEQDIIISLMDQSLKCVKLTPLCVTLNCKOV-NATNTNDSEGTMR 119

Db 90 ENFMNMNNNNVQOEQDIIISLMDQSLKCVKLTPLCVTLNCKOV-NATNTNDSEGTMR 149

Qy 120 GEINCSFNITTSIRDEVQKEXALFYLDVYXIDNNNTSTRLLSCDTSVITQACPKISFE 179

Db 150 GEINCSFNITTSIRDEVQKEXALFYLDVYXIDNNNTSTRLLSCDTSVITQACPKISFE 209

Qy 180 PIPHYCAPAGFALFKCNDTFNGKXPCXNVSTQCTHGIRPVYSTOLLNGSLAESEV 239

Db 210 PIPHYCAPAGFALFKCNDTFNGKXPCXNVSTQCTHGIRPVYSTOLLNGSLAESEV 269

Qy 240 IRSDFNTNAKTIIVQKESVEINCTRPNNNG-----AGDIRQACHN 281

Db 270 IRSDFNTNAKTIIVQKESVEINCTRPNNNG-----AGDIRQACHN 329

Qy 282 ISRAKNDTLKQIYIKRDEQENKTYFNNHSSGGDEPIVHSHFNGGSEFFYCNSTOLFNS 341

Db 330 ISGEKNNNTLKQIYIKRDEQENKTYFNNHSSGGDEPIVHSHFNGGSEFFYCNSTOLFNS 389

Qy 342 TANNTEGSNNTEGNTITLPCRIKQIINMMQEVGKAYAPPPIRGQIRCSSNITGLLTPD 401

Db 390 TW-NNTIGPNNNG-TITLPCRIKQIINMMQEVGKAYAPPPIRGQIRCSSNITGLLTPD 447

Qy 402 GGIN-ENGTEIFRPGGDMRDNWSELYKKVVKIEPLGYAPTCKXRVVQREKRAVGIG 460

Db 448 GKEISNTTEIFRPGGDMRDNWSELYKKVVKIEPLGYAPTCKXRVVQREKRAVGIG 507

Qy 461 AVFLGFIAGAAGSTMGASMTLTVQARLLLSGIYQOQNNLLRAIEAQMLOLTWGIKOL 520

Db 508 AMFLGFIAGAAGSTMGASMTLTVQARLLLSGIYQOQNNLLRAIEAQMLOLTWGIKOL 567

Qy 521 QARVLAVERYLGDQQLIGWCGSKGLICTTAVPNNASMSKSLDRINNNMTWMEEREID 580

Db 568 QARVLAVERYLGDQQLIGWCGSKGLICTTAVPNNASMSKSLDRINNNMTWMEEREID 627

Qy 581 NYTSEIYTLIEESONQOEKNEQELLELDKNASLNNWFDITKMLWY 625

Db 628 NYTSEIYTLIEESONQOEKNEQELLELDKNASLNNWFDITKMLWY 672

RESULT 8

AAB61506

AAB61506 standard; Protein; 579 AA.

AC AAB61506;

DT 05-APR-2001 (first entry)

DE HIV-1 deltaVI2+ SOS gp140 glycoprotein.

XX gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp140; gp140.

Human immunodeficiency virus type 1.

MO200100648-A1.

04-JAN-2001.

23-JUN-2000; 2000WO-US17267.

25-JUN-1999; 99US-0340992.

(PROG-) PROGENICS PHARM INC.

(AARO-) AARON DIAMOND AIDS RES CENT.

Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JF;

WPI; 2001-122993/13.

N-PSDB; AAF28582.

New viral envelope proteins, useful for producing vaccines to treat



```

Db 265 IRSEFTDNAKTIIVOLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAHCN 324
Qy 282 ISRAKMDTLKQIYIKLREOFENKTIYFNHSSGGDEPEIVMHSFNGGSEFFFCNSTOLFNS 341
Db 325 ISGERKMNNTLKQIYTKLQAOFGNKTIYFKOSSGDEPEIVMHSFNGGSEFFFCNSTOLFNS 384
Qy 342 TWNNTEGSNNTGNTITLPCRIKQIIMMOEVEGKAMVAPPRIQIRCSSNITGLLTRD 401
Db 385 TW-NTTIGPNNTNG-TITLPCRIKQIIMMOEVEGKAMVAPPRIQIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKRRVQREKRAVIG 460
Db 443 GGEKISNTTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKRRVQREKRAVIG 502
Qy 461 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIRQL 520
Db 503 AMFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIRQL 562
Qy 521 QARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASMSKSLDRINNMNTMEMEREID 580
Db 563 QARVLAVERYLKDQOLLGIWCGSGKLICTTAVPNNASMSKSLDRINNMNTMEMEREID 622
Qy 581 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKMLWY 625
Db 623 NYTMLIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKMLWY 667

```

## RESULT 10

ABB06211  
ID ABB06211 standard; Protein; 842 AA.

AC ABB06211;

DT 15-MAY-2002 (first entry)

XX HIV Env isolate SF162 amino acid sequence.

XX Human immunodeficiency virus type C; antigenic HIV type C protein;

KM immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;

KM immunostimulant; gene therapy.

OS Human immunodeficiency virus type C.

OS Synthetic.

PN WO200204493-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.

XX 05-JUL-2000; 2000US-0610313.

XX (CHIR) CHIRON CORP.

XX (UYST-) UNIV STELLENBOSCH.

XX Zur Megede J, Barnett SM, Engelbrecht S, Van Rensburg EJ;

XX WPI; 2002-154920/20.

XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful

XX in applications including DNA immunization or generation of packaging

XX cell lines, particularly in gene therapy

XX Disclosure; Fig 105; 233pp; English.

XX The present invention describes expression cassettes comprising a

XX polynucleotide sequence encoding a polypeptide comprising immunogenic

XX HIV type C polypeptides. The expression cassettes comprise any of the

XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or

XX Nef (1). (1) have immunostimulant activity and can be used in gene

XX therapy. The HIV type C polynucleotides are useful in applications

XX including DNA immunisation, generation of packaging cell lines, and

CC production of HIV Type C proteins. The polynucleotides are particularly

CC useful in gene therapy and DNA immunisation applications. ABL39942 to

CC ABL40054 and ABB06215 represent sequences used in the

CC exemplification of the present invention.

XX Sequence 842 AA;

Qy Query Match 87.1%; Score 2918; DB 23; Length 842;

Db Best Local Similarity 85.7%; Pred. No. 3,4e-163;

Qy Matches 553; Conservative 28; Mismatches 42; Indels 22; Gaps 5;

Qy 1 VEKLMVTYVYGVVWPKKATITLFCASDAKADTEVHNWATHACVPTDPNDEVLNVT 60

Db 25 VEKLMVTYVYGVVWPKKATITLFCASDAKADTEVHNWATHACVPTDPNDEVLNVT 84

Qy 61 EHEFMKNNWVEQNEOEDILSLMDQSLKPCVLTPLCYTLNCKOV-NATNTNDEGEMER 119

Db 85 EHEFMKNNWVEQNEOEDILSLMDQSLKPCVLTPLCYTLNCKOV-NATNTNDEGEMER 144

Qy 120 GEIKNCSFNITTSIRDEYKSEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179

Db 145 GEIKNCSFNITTSIRDEYKSEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 204

Qy 180 PIPHYCAPAGFALIKCNKDFNGSKPCXNVSTYQCTHGIRPVVSTOLLNGSLAEEVY 239

Db 205 PIPHYCAPAGFALIKCNKDFNGSKPCXNVSTYQCTHGIRPVVSTOLLNGSLAEEVY 264

Qy 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281

Db 265 IRSEFTDNAKTIIVOLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAHCN 324

Qy 282 ISRAKMDTLKQIYIKLREOFENKTIYFNHSSGGDEPEIVMHSFNGGSEFFFCNSTOLFNS 341

Db 325 ISGERKMNNTLKQIYTKLQAOFGNKTIYFKOSSGDEPEIVMHSFNGGSEFFFCNSTOLFNS 384

Qy 342 TWNNTEGSNNTGNTITLPCRIKQIIMMOEVEGKAMVAPPRIQIRCSSNITGLLTRD 401

Db 385 TW-NTTIGPNNTNG-TITLPCRIKQIIMMOEVEGKAMVAPPRIQIRCSSNITGLLTRD 442

Qy 402 GGIN-ENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKRRVQREKRAVIG 460

Db 443 GGEKISNTTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKRRVQREKRAVIG 502

Qy 461 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIRQL 520

Db 503 AMFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIRQL 562

Qy 521 QARVLAVERYLGDQOLLGIWCGSGKLICTTAVPNNASMSKSLDRINNMNTMEMEREID 580

Db 563 QARVLAVERYLKDQOLLGIWCGSGKLICTTAVPNNASMSKSLDRINNMNTMEMEREID 622

Qy 581 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKMLWY 625

Db 623 NYTMLIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKMLWY 667

## RESULT 11

AAE60422  
ID AAE60422 standard; Protein; 868 AA.

AC AAE60422;

DT 25-MAR-2003 (updated)

DT 20-AUG-1991 (first entry)

XX Sequence of LAV virus env protein.

XX AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.

XX Lymphadenopathy virus.

XX WO8602383-A.

PD 24-APR-1986.  
 XX 18-OCT-1985; 85WO-EP00548.  
 XX 18-OCT-1984; 84GB-0016013.  
 PR 16-NOV-1984; 84GB-0029099.  
 PR 21-JAN-1985; 85GB-0001473.  
 XX (INSP) INST PASTEUR.  
 PA (CNRS) CENT NAT RECH SCIENTIF.  
 XX Montagnier L, Krust B, Chameret F, Chermann JC, Barresinou F;  
 PI Alizon M, Sonigo P;  
 XX WPI: 1986-119166/18.  
 DR N-PSDB; AAN60365.  
 XX  
 PT Purified glyco:protein and peptide(s) - are recognised by sera contg.  
 PT antibodies against lymphadenopathy virus and useful in detecting  
 AIDS antibodies or in vaccines  
 XX  
 XX Disclosure: Fig 4; 75pp; English.  
 CC The inventors claim a polypeptide which is recognised by sera of  
 CC human origin contg. antibodies against the virus of  
 CC lymphadenopathies (LAV) or acquired immune deficiency syndrome  
 CC (AIDS). Also claimed are various peptides corresp. to the AA  
 CC sequences deducible from proteins encoded by LAV DNA, defined by  
 CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance  
 CC with a formula given in the specification.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 868 AA;  
 Query Match 86.5%; Score 2898.5; DB 7; Length 868;  
 Best Local Similarity 84.1%; Pred. No. 4,9e-162;  
 Matches 551; Conservative 32; Mismatches 41; Indels 31; Gaps 7;  
 QY 2 EKLNTVTYGVVWKEATTTTCASDAKAYDTEVNNWATACVPTDPNPOEVLENVTE 61  
 DB 39 EKLNTVTYGVVWKEATTTTCASDAKAYDTEVNNWATACVPTDPNPOEVLENVTE 98  
 QY 62 HNNMMKNWVMOEDTSLMDOSLKPCKLTPCLCTLNCKV-NATNT---TNDSE- 115  
 DB 99 NNNMMKNWVMOEDTSLMDOSLKPCKLTPCLCTLNCKV-NATNT---TNDSE- 158  
 DB 116 -TMRGEIKNCSFNITTSIRDEVOKEVALFYKLDVXINNNTSYRLISCDTSVITQACP 174  
 DB 159 MMEKGEIKNCSFNITTSIRDEVOKEVALFYKLDVXINNNTSYRLISCDTSVITQACP 218  
 QY 175 KISFEPPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTYQCTHGIRPVVSTOLLNGSLA 234  
 DB 219 KVSFEPPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTYQCTHGIRPVVSTOLLNGSLA 278  
 QY 235 EEEVIRSDNFTNNKTIIVOLKESVEINCTRPNNN-----GAGDI 275  
 DB 279 EEEVIRSDNFTNNKTIIVOLKESVEINCTRPNNNTRKIRIQRGPGAFVITGKIGM 338  
 QY 276 ROAHNCISRAKNNDTLKOIVIKLRQF-ENKTIIVFNHSGGDEIYVHSGFNGGEFFYCN 334  
 DB 339 ROAHNCISRAKNNDTLKOIVIKLRQF-ENKTIIVFNHSGGDEIYVHSGFNGGEFFYCN 398  
 QY 335 STOLFNSNTNN--NTEGSSNTEG-NTITLPCRKIQIIMMOEVEGKAMYAPIRQIRCS 390  
 DB 339 STOLFNSNTNN--NTEGSSNTEG-NTITLPCRKIQIIMMOEVEGKAMYAPIRQIRCS 458  
 QY 391 SNITGLLLTRDGINENGTEIFRPGGDMRDWRSLYKVKYKLEPLGVAATPKCRRV 450  
 DB 459 SNITGLLLTRDGINENGTEIFRPGGDMRDWRSLYKVKYKLEPLGVAATPKCRRV 518  
 QY 451 QREKRAVGIGAVFLGFLGAGSTMGASMTLITVQARLLLSGLVQOONNLRAIEAQORWL 510

DB 519 QREKRAVGIGAVFLGFLGAGSTMGASMTLITVQARLLLSGLVQOONNLRAIEAQORWL 578  
 QY 511 QLTWGIQLOARVAVERYLEDQOLGIGWSSGKLTCTTANPWNASWNSKLDRIWNNM 570  
 DB 579 QLTWGIQLOARVAVERYLEDQOLGIGWSSGKLTCTTANPWNASWNSKLDRIWNNM 638  
 QY 571 TMMWEEREIDNTSYRILYTLIESONQOEKNEOELLELDKASLMMWFDITKWLAMY 625  
 DB 639 TMMWEEREIDNTSYRILYTLIESONQOEKNEOELLELDKASLMMWFDITKWLAMY 693  
 RESULT 12  
 AA019389  
 ID AA019389 standard; Protein; 868 AA.  
 XX  
 AC AA019389;  
 XX  
 DT 10-DEC-2002 (first entry)  
 XX  
 DE Lymphadenopathy-associated virus env protein.  
 XX  
 KW Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;  
 KW Lymphadenopathy; AIDS; gag; pol; orf Q; env; orf F; anti-HIV; virucide;  
 XX vaccine.  
 OS Lymphadenopathy-associated virus.  
 PN US6440657-B1.  
 PD 27-AUG-2002.  
 XX  
 PF 06-JAN-2000; 2000US-0478492.  
 XX  
 PR 16-NOV-1984; 84GB-0029099.  
 PR 18-OCT-1985; 85CA-0493377.  
 PR 03-NOV-1992; 92US-0970954.  
 PR 17-OCT-1986; 86US-0920119.  
 PR 01-AUG-1989; 89US-0390499.  
 PR 06-DEC-1990; 90US-0622278.  
 PR 20-AUG-1991; 91US-0747506.  
 PR 01-OCT-1993; 93US-0130565.  
 PR 06-JUN-1995; 95US-0466920.  
 PR 05-DEC-1983; 83US-0558109.  
 PR 28-FEB-1985; 85US-0706562.  
 PR 30-AUG-1985; 85US-0771230.  
 PR 30-AUG-1985; 85US-0771247.  
 PR 30-AUG-1985; 85US-0771248.  
 PA (INSP) INST PASTEUR.  
 PA (CNRS) CENT NAT RECH SCI.  
 XX  
 PI Montagnier L, Krust B, Chameret F, Chermann J;  
 PI Barre-Sinoussi F, Alizon M, Sonigo P, Cole S, Danos O;  
 PI Wain-Hobson S;  
 XX  
 DR WPI: 2002-711525/77.  
 DR N-PSDB; AAL49920.  
 XX  
 PT Novel chemically synthesized Human Immunodeficiency Virus envelope  
 PT protein, useful in vitro diagnostic method for the detection of the  
 PT presence or absence of antibodies that bind to antigens of HIV-1  
 PT retrovirus -  
 XX  
 XX Disclosure: Fig 3; 31pp; English.  
 PS The present invention relates to a chemically synthesised env peptide of  
 CC Human Immunodeficiency Virus (HIV) of less than 150 amino acid residues.  
 CC Such a peptide is useful in an in vitro diagnostic method for the  
 CC detection of the presence or absence of antibodies that bind to antigens  
 CC of HIV-1 retrovirus, by contacting the peptide with a biological fluid  
 CC for a time and under conditions sufficient for the peptide and antibody  
 CC in the biological fluid to form a peptide-antibody complex, detecting the

formation of the peptide-antibody complex by comparing the formation of peptide-antibody complex formation with a control sample, where the formation of peptide-antibody complex is correlated with the presence of antibodies that bind to antigens of HIV-1 retrovirus in the biological sample. It can also be used for preparation of vaccine compositions against AIDS or related syndromes. The present sequence is the lymphadenopathy-associated virus LAV env protein. The LAV virus was isolated from a patient with AIDS.

XX Sequence 868 AA;

Query Match 86.4%; Score 2895.5; DB 23; Length 868;  
Best Local Similarity 84.1%; Pred. No. 7.4e-162;  
Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

QY 2 EKLWVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 61
DB 39 EKLWVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 98
DB 62 HNNMKNMVMQMEDIIISLDOSLKPVCVLTPLCTVLTCKDV-NATNT---TNDSEG- 115
DB 99 NFNMMKNDMVEQMHEDIISLDOSLKPVCVLTPLCTVLTCKDVLTGNATNTSSNTSSGE 158
QY 116 -TMRGEIKNCSFNITTSIRDEQKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
DB 159 MMEKEIKNCSFNITTSIRDEQKEYALFYKLDIIPIDNTTSVLTSCNTSVITQACP 218
QY 175 KISEPPIPIHCAAGAFALIKCNDKTFNGKPCCKNVSTXOCTHGIRPVVSTOLLNGSLA 234
DB 219 KISEPPIPIHCAAGAFALIKCNDKTFNGKPCCKNVSTXOCTHGIRPVVSTOLLNGSLA 278
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 279 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 338
QY 276 ROACHNISRAKNDTLKQIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGEPFYCN 334
DB 339 ROACHNISRAKNDTLKQIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGEPFYCN 398
QY 335 STOLFNSGTWNN---NTBSSNTEG-NTTLPFCRIKQIINMOEYKAMVAPPIGQIRCS 390
DB 399 STOLFNSGTWNN---NTBSSNTEG-NTTLPFCRIKQIINMOEYKAMVAPPIGQIRCS 458
QY 391 SNITGLLTRDGINENGTEIFRPGGDMRDNMWSELYKYVVKIEPLGVAPTKCKRRV 450
DB 459 SNITGLLTRDGINENGTEIFRPGGDMRDNMWSELYKYVVKIEPLGVAPTKCKRRV 518
QY 451 OREKRAVIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQORML 510
DB 519 OREKRAVIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQORML 578
QY 511 QLTWGIQOLQARVLAVERLYGDDQOLIGWSSGGLICTAVPNNASNSKSLDINNM 570
DB 579 QLTWGIQOLQARVLAVERLYGDDQOLIGWSSGGLICTAVPNNASNSKSLDINNM 638
QY 571 TWMEEREIDNVTSEIYTLIESNOQEKNEOELLELDKASLMMFDTTKLMT 625
DB 639 TWMEEREIDNVTSEIYTLIESNOQEKNEOELLELDKASLMMFDTTKLMT 693

```

RESULT 13  
AAP60063  
ID AAP60063 standard; Protein; 868 AA.

XX AAP60063;  
XX 25-MAR-2003 (updated)  
XX 31-OCT-2002 (updated)  
XX 11-JUN-1991 (first entry)  
XX HIV virus env gene product.  
XX DE  
XX HIV virus; vaccine; AIDS.

XX OS Human immunodeficiency virus.

XX EN EP201716-A.

XX PD 20-NOV-1986.

XX PF 07-APR-1986; 86EP-0104741.

XX PR 07-NOV-1985; 85US-0783229.

XX PR 08-APR-1985; 85US-0721237.

XX PR 29-APR-1985; 85US-0728052.

XX PR 10-JUN-1985; 85US-0753769.

XX PA (GENE-) GENETIC SYSTEMS CORP.

XX PA (BRIM) BRISTOL MYERS SQUIBB CO.

XX PA Watanabe SM, Cosand W, Mardale S, Travis BW;

XX PI WPI; 1986-306545/47.

XX DR N-PSDB; AAN60077.

XX PT New DNA sequence comprising part of env. region of LAV genome -

XX PT and corresp. transformational protein prod., useful in vaccine

XX PT against AIDS.

XX PS Disclosure; Fig 8; 27pp; English.

XX CC The protein is immunologically reactive with antibodies against

XX CC HIV virus, and can be used in a vaccine composition for

XX CC prevention against AIDS.

XX CC (Updated on 31-OCT-2002 to add missing OS field.)

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 868 AA;

Query Match 86.4%; Score 2894.5; DB 7; Length 868;

Best Local Similarity 84.1%; Pred. No. 8.4e-162;

Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

QY 2 EKLWVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 61
DB 39 EKLWVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 98
DB 62 HNNMKNMVMQMEDIIISLDOSLKPVCVLTPLCTVLTCKDV-NATNT---TNDSEG- 115
DB 99 NFNMMKNDMVEQMHEDIISLDOSLKPVCVLTPLCTVLTCKDVLTGNATNTSSNTSSGE 158
QY 116 -TMRGEIKNCSFNITTSIRDEQKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
DB 159 MMEKEIKNCSFNITTSIRDEQKEYALFYKLDIIPIDNTTSVLTSCNTSVITQACP 218
QY 175 KISEPPIPIHCAAGAFALIKCNDKTFNGKPCCKNVSTXOCTHGIRPVVSTOLLNGSLA 234
DB 219 KISEPPIPIHCAAGAFALIKCNDKTFNGKPCCKNVSTXOCTHGIRPVVSTOLLNGSLA 278
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 279 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 338
QY 276 ROACHNISRAKNDTLKQIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGEPFYCN 334
DB 339 ROACHNISRAKNDTLKQIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGEPFYCN 398
QY 335 STOLFNSGTWNN---NTBSSNTEG-NTTLPFCRIKQIINMOEYKAMVAPPIGQIRCS 390
DB 399 STOLFNSGTWNN---NTBSSNTEG-NTTLPFCRIKQIINMOEYKAMVAPPIGQIRCS 458
QY 391 SNITGLLTRDGINENGTEIFRPGGDMRDNMWSELYKYVVKIEPLGVAPTKCKRRV 450
DB 459 SNITGLLTRDGINENGTEIFRPGGDMRDNMWSELYKYVVKIEPLGVAPTKCKRRV 518
QY 451 OREKRAVIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQORML 510

```

```

Db      519 OREKAVGALFLFELGAGSTMGARSMTLLTVQARQLSGIVQOQNNLRAIEAQHLL 578
Qy      511 QLTWGIKQLOARVLAVERYLGDQQLIGWCSGKLICTTAVPMNASMSKSLDRIMNMN 570
Db      579 QLTWGIKQLOARVLAVERYLKDQQLIGWCSGKLICTTAVPMNASMSKSLDRIMNMN 638
Qy      571 TWMEREIDNNTSYTYLIESONQOEKNEQBELLDKMSLMMWFDTIKMLMY 625
Db      639 TWMEREIDNNTSYTYLIESONQOEKNEQBELLDKMSLMMWFDTIKMLMY 693

RESULT 14
AAB85999 standard; Protein; 856 AA.
AC      AAB85999;
DT      30-NOV-2001 (first entry)

Amino acid sequence of HIV-1 isolate BH10 envelope polypeptide.
KM      HIV-1; gp120; BH10; vaccine; immunization; envelope polypeptide.
XX      Human immunodeficiency virus type 1.
XX      Key Location/Qualifiers
XX      FT Peptide 79..184
XX      FT Peptide /note="specifically claimed fragment"
XX      FT Peptide 326..400
XX      FT Peptide /note="specifically claimed fragment"

US6268484-B1.
XX      31-JUL-2001.
XX      30-JUL-1998; 98US-0124900.
XX      07-JUN-1995; 95US-0478536.
XX      19-APR-1995; 95WO-EP01481.
XX      (POLY-) POLY(MN SCI IMMUNOBIOLOGISCHE FORSCHUNG.
XX      Katinger H, Buchacher A, Ernst W, Ballaun C, Purtscher M;
XX      Trkola A, Predl R, Schmatz C, Klima A, Steindl F, Muster T;
XX      WPI; 2001-556601/62.
XX      N-PSDB; AAH76385.

New anti-idiotypic antibodies consisting of one or both amino acid
PT      sequences corresponding to amino acid positions 79-84 or 326-400 of the
PT      processed gp120 of HIV-isolate BH10, useful as vaccine against HIV-1
PT      infections.
XX      Claim 1; Columns 33-38; 27pp; English.
XX      The invention relates to a peptide fragment comprising of amino acid
XX      sequences corresponding to sequences within the processed gp120 of HIV-1
XX      isolate BH10 (Genbank accession M15654). The peptides are useful in the
XX      detection, prevention and treatment of HIV-1 infections, and in AIDS
XX      therapy. The antibodies are especially useful as vaccines for active and
XX      passive immunization, or for the detection and/or determination of HIV-1
XX      infected cells and/or HIV-1 viruses. The present sequence represents the
XX      amino acid sequence of a envelope polypeptide of HIV-1 isolate BH10
XX      (Genbank accession M15654).
XX      Sequence 856 AA;
SQ      Query Match 86.2%; Score 2887; DB 22; Length 856;
        Best Local Similarity 84.0%; Pred. No. 2.3e-161;
        Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;
Qy      2 EKLMTVTVYGVVWKEATTTLFCASDAKAYDTEVHNWMTACVPTDPNPQEVVLNVT 61

```

```

Db      32 EKLMTVTVYGVVWKEATTTLFCASDAKAYDTEVHNWMTACVPTDPNPQEVVLNVT 91
Qy      62 HNNMKNMNMVEMOMDIIISMDQSLKPCVKLTPLCVTLNCKOVNMTNTNDEG--TMR 119
Db      92 NNNMKNMNMVEMOMDIIISMDQSLKPCVKLTPLCVTLNCKOVNMTNTNDEG--TMR 151
Qy      120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVAIDNNNTSRYLISCDTSYTOACPISPE 179
Db      152 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVAIDNNNTSRYLISCDTSYTOACPISPE 211
Qy      180 PIPHYCAPAGFAILKCKDTFNGKPCKNVSTXOCTGIRPVSTOLLNGLAEEBV 239
Db      212 PIPHYCAPAGFAILKCKDTFNGKPCKNVSTXOCTGIRPVSTOLLNGLAEEBV 271
Qy      240 IRSDFNTNNAKTIYQLKESVEINCTRNPN-----GAGDIRQAMC 280
Db      272 IRSDFNTNNAKTIYQLKESVEINCTRNPN-----GAGDIRQAMC 331
Qy      281 NISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSGDPEIVMHSFNGGEFFYCNSSTOLF 339
Db      332 NISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSGDPEIVMHSFNGGEFFYCNSSTOLF 391
Qy      340 NSTWNN--NTEGSNNTEG-NTITLPCRIKOIINMGEVKAHYAPPIRGQIRCSNITG 395
Db      392 NSTWNN--NTEGSNNTEG-NTITLPCRIKOIINMGEVKAHYAPPIRGQIRCSNITG 451
Qy      396 LLITRDGGINENGTIFRPGGDMNDNRSELYKRVYKIEPLGVAPTCKRRVYQREKR 455
Db      452 LLITRDGGINENGTIFRPGGDMNDNRSELYKRVYKIEPLGVAPTCKRRVYQREKR 511
Qy      456 AVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAQHLLQLTW 515
Db      512 AVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAQHLLQLTW 571
Qy      516 GIKQLOARVLAVERYLGDQQLIGWCSGKLICTTAVPMNASMSKSLDRIMNMN 575
Db      572 GIKQLOARVLAVERYLGDQQLIGWCSGKLICTTAVPMNASMSKSLDRIMNMN 631
Qy      576 EREINDNNTSYTYLIESONQOEKNEQBELLDKMSLMMWFDTIKMLMY 625
Db      632 EREINDNNTSYTYLIESONQOEKNEQBELLDKMSLMMWFDTIKMLMY 681

RESULT 15
AAR43869 standard; Protein; 863 AA.
AC      AAR43869;
DT      25-MAR-2003 (updated)
DT      23-DEC-1993 (first entry)
XX      HTLV-III ENV-LOR gene product (Clone BH10).
XX      Polypeptide; antibodies; HTLV; AIDS; vaccine.
XX      Human T-cell lymphotropic virus.
XX      EP552850-A1.
XX      28-JUL-1993.
XX      10-OCT-1985; 93EP-0200929.
XX      10-OCT-1984; 84US-0659339.
XX      23-JAN-1985; 85US-0693866.
XX      10-OCT-1985; 85EP-0307260.
XX      (GENZ ) CENTOCOR INC.
XX      Chang NT, Gallo RC, Wong-staal F;
XX

```

DR WPI: 1993-236543/30.  
DR N-PSDB: AAQ45919.

XX Cloning and expression of new HTLV-III DNA - used to obtain  
PT polypeptide(s) and antibodies for diagnosis, prevention and  
PT treatment of HTLV-III infection, partic. AIDS

XX Disclosure; Figure 3; 31pp; English.

CC A fragment of DNA approximately 200-500 base pairs in length is  
CC ligated into a vector and used to transform E.coli. These cells then  
CC express a polypeptide which is immunoreactive with HTLV-III-specific  
CC antibody. The HTLV-III polypeptides can be used for the production  
CC of antibodies, in immunoassays for the detection of HTLV-  
CC III-specific antibodies and in vaccines for the prevention of AIDS.  
CC The antibodies can also be used to detect HTLV-III polypeptides.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
CC (Updated on 25-MAR-2003 to correct PR field.)  
CC (Updated on 25-MAR-2003 to correct PR field.)

Sequence 863 AA:

Query Match 86.2%; Score 2887; DB 14; Length 863;  
Best Local Similarity 84.0%; Pred. No. 2.3e-161;  
Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

```
QY 2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 61
DB 39 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 98
QY 62 HFNWKNMVEOMOEEDISLWDOSLPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 99 NFNWKNMVEOMOEEDISLWDOSLPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 158
QY 120 GEIKKCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISCTSVITQACPXISFE 179
DB 159 GEIKKCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISCTSVITQACPXISFE 218
QY 180 PIPHYCAPAGPAILKCKDKTFNGKXPKCNVSTXCTGIRPVVSTOLLNGSLAEEV 239
DB 219 PIPHYCAPAGPAILKCKDKTFNGKXPKCNVSTXCTGIRPVVSTOLLNGSLAEEV 278
QY 240 IRSDFNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAH 280
DB 279 IRSANFTDNAKTIIVQLNQSVEINCTRPNNRKSIRIQRGRAFYTIIGKIGNRQAH 338
QY 281 NISRAKMDTLKQIVIKRBOF-ENKTI VFNHSSGDPEI VMSFNGCGEPFYCNSTOLF 339
DB 339 NISRAKMDTLKQIVIKRBOF-ENKTI VFNHSSGDPEI VMSFNGCGEPFYCNSTOLF 398
QY 340 NSTWNN--NTEGNNTEG-NTITLPCRIKQIINNMOEVGKAMVAPRIGQIRCSNITG 395
DB 399 NSTWNNSTWSTKSGNNTGSDITLPCRIKQIINNMOEVGKAMVAPRIGQIRCSNITG 458
QY 396 LLTTRDGINENGTETFRPGGDMRDNMRSELYKVKVIEPLGVAPTCKRRVVQREKR 455
DB 459 LLTTRDGINENGTETFRPGGDMRDNMRSELYKVKVIEPLGVAPTCKRRVVQREKR 518
QY 456 AVGIGAVFLGPGAAGSTMGASMTLTVQARILLSGIVQOQNNLLRAIEAQRMQLTVM 515
DB 519 AVGIGAVFLGPGAAGSTMGASMTLTVQARILLSGIVQOQNNLLRAIEAQRMQLTVM 578
QY 516 GIKOLQAVLAVERYLQOQLIGTGGCGKLICTTAVPMNASSWSKSLDRIINNNTMMEW 575
DB 579 GIKOLQAVLAVERYLQOQLIGTGGCGKLICTTAVPMNASSWSKSLDRIINNNTMMEW 638
QY 576 EREIDNYSSEIYTLIESQNOQEKNEOEILBLDKWASLMMWFDITKMLY 625
DB 639 DREINNYSLIHSLEBSQNOQEKNEOEILBLDKWASLMMWFDITKMLY 688
```

Search completed: December 12, 2003, 12:30:07  
Job time : 39.2225 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:28:04 ; Search time 14.2122 Seconds  
(without alignments)  
1860.670 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351

Sequence: 1 VEKLWTVYVGVPMKEATT.....ELDKWASLWNFDTKWLWY 625

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/aa/5A\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/aa/6A\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/aa/6A\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/aa/6A\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2968.5	88.6	850	2	US-08-448-603A-28
2	2968.5	88.6	850	3	US-09-134-075-28
3	2968.5	88.6	850	4	US-09-492-739-28
4	2938.5	87.7	855	1	US-08-022-835-6
5	2938.5	87.7	855	1	US-08-388-809-6
6	2938.5	87.7	855	2	US-08-647-714-6
7	2903.5	86.6	861	1	US-07-956-483-14
8	2901.5	86.6	861	1	US-08-127-499A-14
9	2901.5	86.6	861	3	US-08-482-847-14
10	2901.5	86.6	861	3	US-07-956-483-10
11	2901.5	86.6	861	3	US-08-472-240A-1
12	2901.5	86.6	861	3	US-08-472-240A-7
13	2901.5	86.6	861	4	US-08-817-441-103
14	2893.5	86.2	861	3	US-07-956-483-16
15	2887	86.2	863	3	US-09-124-800-9
16	2887	86.2	863	3	US-08-463-210-11
17	2886	86.1	826	1	US-08-375-510-2
18	2886	86.1	826	2	US-08-487-657-2
19	2886	86.1	826	4	US-09-309-572-23
20	2878	85.9	856	2	US-07-916-098A-2
21	2876	85.8	857	3	US-08-448-603A-30
22	2876	85.8	857	2	US-09-134-075-30
23	2876	85.8	857	4	US-09-492-739-30
24	2872.5	85.7	839	4	US-08-472-240A-10
25	2870	85.6	856	1	US-08-022-835-2
26	2870	85.6	856	1	US-08-388-809-2
27	2870	85.6	856	2	US-08-647-714-2

28	2870	85.6	856	3	US-07-956-483-11	Sequence 11, Appl
29	2867	85.6	856	4	US-09-337-387-11	Sequence 11, Appl
30	2861	85.4	857	1	US-08-022-835-4	Sequence 4, Appl
31	2861	85.4	857	1	US-08-388-809-4	Sequence 4, Appl
32	2861	85.4	857	2	US-08-647-714-4	Sequence 4, Appl
33	2854.5	85.2	880	2	US-08-788-815-7	Sequence 7, Appl
34	2854.5	85.2	880	3	US-09-157-963-7	Sequence 7, Appl
35	2852	85.1	887	3	US-08-472-240A-5	Sequence 5, Appl
36	2828	84.4	855	3	US-07-956-483-15	Sequence 15, Appl
37	2822.5	84.2	726	4	US-09-337-387-3	Sequence 3, Appl
38	2817	84.1	856	1	US-08-375-100-1	Sequence 1, Appl
39	2817	84.1	855	3	US-07-956-483-13	Sequence 13, Appl
40	2815.5	84.0	759	4	US-09-337-387-12	Sequence 12, Appl
41	2809	83.8	887	3	US-08-472-240A-2	Sequence 2, Appl
42	2797	83.5	887	3	US-08-472-240A-6	Sequence 6, Appl
43	2788	83.2	887	2	US-08-472-240A-4	Sequence 4, Appl
44	2760.5	82.4	665	2	US-08-448-603A-32	Sequence 32, Appl
45	2760.5	82.4	665	3	US-09-134-075-32	Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-08-448-603A-28  
; Sequence 28, Application US/08448603A  
; Patent No. 5864027  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McClutchen, Doyle, Brown & Eriksen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,603A  
; FILING DATE: 07-JUN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/072,833  
; FILING DATE: 07-JUN-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halliday, Emily  
; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 850 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-448-603A-28

Query Match 88.6%; Score 2968.5; DB 2: Length 850;  
Best Local Similarity 86.9%; Pred. No. 8.8e-248;  
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;  
2 EKLWTVYVGVPMKEATTTTFCASDAKAVDEVHNVMAVTHACVPTDPNQEVLNATVE 61  
|||||

```

Db      31 EKLWTVYVGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEIGLENTVE 90
Qy      62 HNNMKNMWEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERG 120
Db      91 NFNMMKNMWEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERG 150
Qy      121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151 EIKKCSFNITTSIRKMKNEVALFYKLDVPIIDNNTSYRLISCDTSVITQACPKISFEP 210
Qy      181 IPIHYCAPAGAILKCDNKTNGKPCKNVSTXOCTHGIRPVYSTOLLNGLSLAEVEVI 240
Db      211 IPIHYCAPAGAILKCDNKTNGKPCKNVSTXOCTHGIRPVYSTOLLNGLSLAEVEVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHONI 282
Db      271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
Qy      283 SRAKMNNTLKOIIVKLAEOPEFNKTIIVFNHSSGDEPEIVMHSFNGGFEFCNSTQLFNST 342
Db      331 SSTMNNNTLKOIIVKLAEOPEFNKTIIVFNHSSGDEPEIVMHSFNGGFEFCNSTQLFNST 389
Qy      343 WN-----NTEGSNNTBENTITLPCRIOIIMMOEYKAMVAPPIRGQIRCSNITGL 397
Db      390 WNYTYTWNNTBGSNDT-GRNITLQCRIOIIMMOEYKAMVAPPIRGQIRCSNITGL 448
Qy      398 LTRDGGINENGEIIFRPGGDMRDNRSELKYKVKIIEPLGVAPTYKRRVVOREKRAV 457
Db      449 LTRDGG-NNSETEIFRPGGDMRDNRSELKYKVKIIEPLGVAPTYKRRVVOREKRAV 507
Qy      458 GIGAVFLGFLGAAGSTWGAASVTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTWGI 517
Db      508 GIGAVFLGFLGAAGSTWGAASVTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTWGI 567
Qy      518 KOLQARVLAVERYLKDQOLLGIWCSGKLICTTAVPNAASWSNKSLDRIINNMTWEMER 577
Db      568 KOLQARVLAVERYLKDQOLLGIWCSGKLICTTAVPNAASWSNKSLDRIINNMTWEMER 627
Qy      578 EIDNVTSEIYTLIESONQOEKNEQELLELDKWSLNNWPDITKWLWY 625
Db      628 EIDNVTSEIYTLIESONQOEKNEQELLELDKWSLNNWPDITKWLWY 675

RESULT 2
US-09-134-075-28
; Sequence 28, Application US/09134075
; Patent No. 6042836
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
APPLICANT: Nakamura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: McCutchen, Doyle, Brown & Enersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASCSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134.075
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/448,603
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily

```

```

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEFAX: 415-393-2286
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-134-075-28

Query Match      88.6%; Score 2968.5; DB 3; Length 850;
Best Local Similarity 86.9%; Pred. No. 8.8e-248;
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

Qy      2 EKLWTVYVGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEIGLENTVE 61
Db      31 EKLWTVYVGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEIGLENTVE 90
Qy      62 HNNMKNMWEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERG 120
Db      91 NFNMMKNMWEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMERG 150
Qy      121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151 EIKKCSFNITTSIRKMKNEVALFYKLDVPIIDNNTSYRLISCDTSVITQACPKISFEP 210
Qy      181 IPIHYCAPAGAILKCDNKTNGKPCKNVSTXOCTHGIRPVYSTOLLNGLSLAEVEVI 240
Db      211 IPIHYCAPAGAILKCDNKTNGKPCKNVSTXOCTHGIRPVYSTOLLNGLSLAEVEVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHONI 282
Db      271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
Qy      283 SRAKMNNTLKOIIVKLAEOPEFNKTIIVFNHSSGDEPEIVMHSFNGGFEFCNSTQLFNST 342
Db      331 SSTMNNNTLKOIIVKLAEOPEFNKTIIVFNHSSGDEPEIVMHSFNGGFEFCNSTQLFNST 389
Qy      343 WN-----NTEGSNNTBENTITLPCRIOIIMMOEYKAMVAPPIRGQIRCSNITGL 397
Db      390 WNYTYTWNNTBGSNDT-GRNITLQCRIOIIMMOEYKAMVAPPIRGQIRCSNITGL 448
Qy      398 LTRDGGINENGEIIFRPGGDMRDNRSELKYKVKIIEPLGVAPTYKRRVVOREKRAV 457
Db      449 LTRDGG-NNSETEIFRPGGDMRDNRSELKYKVKIIEPLGVAPTYKRRVVOREKRAV 507
Qy      458 GIGAVFLGFLGAAGSTWGAASVTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTWGI 517
Db      508 GIGAVFLGFLGAAGSTWGAASVTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTWGI 567
Qy      518 KOLQARVLAVERYLKDQOLLGIWCSGKLICTTAVPNAASWSNKSLDRIINNMTWEMER 577
Db      568 KOLQARVLAVERYLKDQOLLGIWCSGKLICTTAVPNAASWSNKSLDRIINNMTWEMER 627
Qy      578 EIDNVTSEIYTLIESONQOEKNEQELLELDKWSLNNWPDITKWLWY 625
Db      628 EIDNVTSEIYTLIESONQOEKNEQELLELDKWSLNNWPDITKWLWY 675

RESULT 3
US-09-492-739-28
; Sequence 28, Application US/09492739
; Patent No. 6331404
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
APPLICANT: Nakamura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: McCutchen, Doyle, Brown & Eversen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-492-739-28

Query Match      88.6%; Score 2968.5; DB 4; Length 850;
Best Local Similarity 86.9%; Pred. No.: 8.8e-248;
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

QY 2 EKLMTVVYGVVPMKEATTTLCASADAKAYDTEVHNWATNACVPTDPPQEVLENYTE 61
DB 31 EKLMTVVYGVVPMKEATTTLCASADAKAYDTEVHNWATNACVPTDPPQEVLENYTE 90
QY 62 HNNMKNMNVQEMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSGTMERG 120
DB 91 HNNMKNMNVQEMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSGTMERG 150
QY 121 EIKNCSFNITTSIRDEVOKEVALPYRLDVVXIINNNTSYRLISCDTSVITTOACPKISFEP 180
DB 151 EIKNCSFNITTSIRDMKMEVALPYRLDVVPIIDNDNTSYRLISCDTSVITTOACPKISFEP 210
QY 181 IPIHICAPAGFAILKNDKTFNGKAPCKAVSTYQCTHGIRPVYSTOLLNGSLAEVEEVI 240
DB 211 IPIHICAPAGFAILKNDKTFNGKAPCKAVSTYQCTHGIRPVYSTOLLNGSLAEVEEVI 270
QY 241 REDNFTNNAKTIIVOLKESVVEINCTRPNNNG-----AGDIRAHNCVI 282
DB 271 KSAHNSDNKTIIVOLKESVVEINCTRPNNNTRRSIHIGRAFYATGEIIGDIRAHNCVI 330
QY 283 SHAKNDTLKQIVIKLREOPENKTIIVFNHSSGGDEIIVNHSFNGCGEFYCNSTOLFENST 342
DB 331 SSKKNNTLTKQIVTKLREHF-NKTIIVFNHSSGGDEIIVNHSFNGCGEFYCNSTOLFENST 389
QY 343 WY-----NNTSGSNTTEGTTITLPCRIKQIIMMVGAVGAAAPPIRGQIRGSSNTGLL 397
DB 390 WNYTYTWNTEGSSNTI-GRNITLQCIKQIIMMVGAVGAAAPPIRGQIRGSSNTGLL 448
QY 398 LTRDGINENGTEIFRPGGDMRDNRSELYKKYKVVKIEPLGAPAPKCKRKYVOREKRAV 457
DB 449 LTRDGG-ANSETEIFRPGGDMRDNRSELYKKYKVVKIEPLGAPAPKCKRKYVOREKRAV 507
QY 458 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQQQNNLIRALEAOQRMQLQTVWCI 517
DB 517 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQQQNNLIRALEAOQRMQLQTVWCI 517

```

```

DB 508 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQQQNNLIRALEAOQRMQLQTVWCI 567
QY 518 KOLARVLAVERYLDGDOOLGIGWGSGLICTTAVPAPNANSGNSKSDRIANNMTWMEER 577
DB 568 KOLARVLAVERYLDGDOOLGIGWGSGLICTTAVPAPNANSGNSKSDRIANNMTWMEER 627
QY 578 EIDNVTSEIYTLIBESONQOEKNBOELLELDKMSLWNNFDITKWLWY 625
DB 628 EIDNVTSLIYSLIBESONQOEKNBOELLELDKMSLWNNFDITKWLWY 675

RESULT 4
US-08-022-835-6
Sequence 6, Application US/08022835
Patent No. 5420030
GENERAL INFORMATION:
APPLICANT: Reitz Jr., Marvin S.
APPLICANT: Franchini, Genevieve
APPLICANT: Markham, Phillip D.
APPLICANT: Gallo, Robert C.
APPLICANT: Lori, Franco C.
APPLICANT: Popovic, Mikulas
APPLICANT: Gartner, Suzanne
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
NUMBER OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: Eleventh Floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-022-835-6

Query Match      87.7%; Score 2938.5; DB 1; Length 855;
Best Local Similarity 85.1%; Pred. No.: 3.4e-245;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

QY 2 EKLMTVVYGVVPMKEATTTLCASADAKAYDTEVHNWATNACVPTDPPQEVLENYTE 61
DB 31 EKLMTVVYGVVPMKEATTTLCASADAKAYDTEVHNWATNACVPTDPPQEVLENYTE 90
QY 62 HNNMKNMNVQEMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115
DB 91 HNNMKNMNVQEMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150

```

```

QY 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
D 151 MVGGEMKNCSPNITTNIRGVQKEYALFYKLDIAPIDNNSNNRYRLISCDTSVITQACP 210
QY 175 KISFEPIPIHYCAPAGFAILKCDKTKPKCKPCKNVSTXOCTGIRPVSTQOLLNLSLA 234
D 211 KVSFEPIPIHYCAPAGFAILKCDKTKPKCKPCKNVSTXOCTGIRPVSTQOLLNLSLA 270
QY 235 EEEVVISDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
D 271 EEEVVISANPADNAKTIIVOLNESVEINCTRPNNNRKSIHIGPRAFTTGTGIDIR 330
QY 277 QAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNST 336
D 331 QAHCNLSRAKNDTLNKIVIKLREOFENKTIIVFKHSSGDEPEIVMHSFNCGEFFYCNST 390
QY 337 QLFNSTNNNTGSGNTEGNTITLPCRKOIINMOEYGVKAMVAPPIRGQIRCSNITGL 396
D 391 QLFNSTNNNTGSGNTEGNTITLPCRKOIINMOEYGVKAMVAPPIRGQIRCSNITGL 450
QY 397 LITRDGGINENGTEIFRPGGDMRDNMRSELYKKVYKIEPLGVAFTCKRVRVQREKRA 456
D 451 LITRDGGEDKTEIFRPGGDMRDNMRSELYKKVYKIEPLGVAFTCKRVRVQREKRA 510
QY 457 VGIGAVFLGFLGAAGSTMGAAAMTLTVQARLLSGIVQOONNLLRAIEAQOQLTLTWG 516
D 511 VGIGAVFLGFLGAAGSTMGAAAMTLTVQARLLSGIVQOONNLLRAIEAQOQLTLTWG 570
QY 517 IKQIARVLAVERLYGQOQLIGWCGSKLICCTAVPWNASWSKSLDRINNTMMEWE 576
D 571 IKQIARVLAVERLYGQOQLIGWCGSKLICCTAVPWNASWSKSLDRINNTMMEWE 630
QY 577 REINNTSEIYTLIEESONOEKNEOELELDKMASLNNMFDTIKMLMY 625
D 631 REINNTSIYTLIEESONOEKNEOELELDKMASLNNMFDTIKMLMY 679

RESULT 5
US-08-388-809-6
; Sequence 6, Application US/08388809
; Patent No. 5576000
; GENERAL INFORMATION:
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
; APPLICANT: GENOVEFPA; MARKHAM, PHILIP D. GALLO, ROBERT
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
; APPLICANT: GARTNER, SUZANNE
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK, 3.5"
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,809
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/022,835
; FILING DATE: 25-FEB-1993
; APPLICATION NUMBER: US 07/599,491
; FILING DATE: 17-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: LESLIE A. SERJUNIAN
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4092US2

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-388-809-6

Query Match 87.7%; Score 2938.5; DB 1; Length 855;
Best Local Similarity 85.1%; Pred. No. 3,4e-245;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

QY 2 EKLMTVYGVVWPKKETTTLTFCASDAKAYDEVYHNNMAHACVPTDPNQEVLNVTE 61
D 31 EKLMTVYGVVWPKKETTTLTFCASDAKAYDEVYHNNMAHACVPTDPNQEVLNVTE 90
QY 62 HFNWKNMVEOMQEDIIISLDOSLKPCVKLTPLCVTLNCKDY-----NATNTNDSG 115
D 91 HFNWKNMVEOMQEDIIISLDOSLKPCVKLTPLCVTLNCKDY-----NATNTNDSG 150
QY 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
D 151 MVGGEMKNCSPNITTNIRGVQKEYALFYKLDIAPIDNNSNNRYRLISCDTSVITQACP 210
QY 175 KISFEPIPIHYCAPAGFAILKCDKTKPKCKPCKNVSTXOCTGIRPVSTQOLLNLSLA 234
D 211 KVSFEPIPIHYCAPAGFAILKCDKTKPKCKPCKNVSTXOCTGIRPVSTQOLLNLSLA 270
QY 235 EEEVVISDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
D 271 EEEVVISANPADNAKTIIVOLNESVEINCTRPNNNRKSIHIGPRAFTTGTGIDIR 330
QY 277 QAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNST 336
D 331 QAHCNLSRAKNDTLNKIVIKLREOFENKTIIVFKHSSGDEPEIVMHSFNCGEFFYCNST 390
QY 337 QLFNSTNNNTGSGNTEGNTITLPCRKOIINMOEYGVKAMVAPPIRGQIRCSNITGL 396
D 391 QLFNSTNNNTGSGNTEGNTITLPCRKOIINMOEYGVKAMVAPPIRGQIRCSNITGL 450
QY 397 LITRDGGINENGTEIFRPGGDMRDNMRSELYKKVYKIEPLGVAFTCKRVRVQREKRA 456
D 451 LITRDGGEDKTEIFRPGGDMRDNMRSELYKKVYKIEPLGVAFTCKRVRVQREKRA 510
QY 457 VGIGAVFLGFLGAAGSTMGAAAMTLTVQARLLSGIVQOONNLLRAIEAQOQLTLTWG 516
D 511 VGIGAVFLGFLGAAGSTMGAAAMTLTVQARLLSGIVQOONNLLRAIEAQOQLTLTWG 570
QY 517 IKQIARVLAVERLYGQOQLIGWCGSKLICCTAVPWNASWSKSLDRINNTMMEWE 576
D 571 IKQIARVLAVERLYGQOQLIGWCGSKLICCTAVPWNASWSKSLDRINNTMMEWE 630
QY 577 REINNTSEIYTLIEESONOEKNEOELELDKMASLNNMFDTIKMLMY 625
D 631 REINNTSIYTLIEESONOEKNEOELELDKMASLNNMFDTIKMLMY 679

RESULT 6
US-08-647-714-6
; Sequence 6, Application US/08647714
; Patent No. 5869313
; GENERAL INFORMATION:
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
; APPLICANT: GENOVEFPA; MARKHAM, PHILIP D. GALLO, ROBERT
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
; APPLICANT: GARTNER, SUZANNE
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
; NUMBER OF SEQUENCES: 6

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK, 3.5"
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,714
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/388,809
FILING DATE: 15-FEB-1995
APPLICATION NUMBER: US 08/022,835
FILING DATE: 25-FEB-1993
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: LESLIE A. SERUNIAN
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4092US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-647-714-6

Query Match      87.7% Score 2938.5; DB 2; Length 855;
Best Local Similarity 85.1%; Pred. No. 3.4e-245;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

QY 2 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 115
DB 91 HNNMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 150
QY 116 TWBGEIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSVRLISCDTSVTOACP 174
DB 151 WGGGBMKKCSFNITTNIRKQVKEVALFYKLDIADIDNNSNRRLISCDTSVTOACP 210
QY 175 KISFEPIPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTGIRPVVSTOLLNGSLA 234
DB 211 KVSFEPIPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTGIRPVVSTOLLNGSLA 270
QY 235 EEEVVIRSNFTNNAKTIIIVOLKESYEINCTRPNNNG-----AGDIR 276
DB 271 EEEVVIRSNFTNNAKTIIIVOLKESYEINCTRPNNNG-----AGDIR 330
QY 277 QAHCHISRAKNDTKQOIVIKLREOPENKTIVFNHSSGGDEPVIHMSFNGGSEPFYCNST 336
DB 331 QAHCHISRAKNDTKQOIVIKLREOPENKTIVFNHSSGGDEPVIHMSFNGGSEPFYCNST 390
QY 337 QLFNSTWNNNTGSSNNTGNTTTLPCRKOIIMNQEVKAMVAPPIRQOIRCSSNITGL 396
DB 391 QLFNSTWNNNTGSSNNTGNTTTLPCRKOIIMNQEVKAMVAPPIRQOIRCSSNITGL 450
QY 397 LLTRGGINGENGTLEFRGGGDMRDWMBELYYKKVYKEPIGVAPTCKRRVYOREKKA 456
DB 451 LLTRGGINGENGTLEFRGGGDMRDWMBELYYKKVYKEPIGVAPTCKRRVYOREKKA 510

```

```

QY 457 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMLOLTWVG 516
DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMLOLTWVG 570
QY 517 IKOLARVLAVERYIGDQQLGIGWCSGKLTCTVAPNNAWSNKSLEDRINNMWMEWE 576
DB 571 IKOLARVLAVERYIGDQQLGIGWCSGKLTCTVAPNNAWSNKSLEDRINNMWMEWE 630
QY 577 REINDYSEIYTLIESONQOEKNOEPLLEIDKASLNNWPDITKMWY 625
DB 631 REINDYSEIYTLIESONQOEKNOEPLLEIDKASLNNWPDITKMWY 679

RESULT 7
US-07-956-483-14
Sequence 14, Application US/07956483
Patent No. 6261799
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,483
FILING DATE: 31-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/19742
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 05392
FILING DATE: 02-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feuilly, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 01753-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match      86.6% Score 2903.5; DB 3; Length 855;
Best Local Similarity 84.4%; Pred. No. 3.6e-242;
Matches 553; Conservative 31; Mismatches 36; Indels 35; Gaps 8;

QY 2 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 115
DB 91 HNNMKNMNVEMQMDIISLMDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 150

```

```

QY 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 175
DB 151 KMEGEMTNCSFNITTSIRSKQKEYALFYKLDVVPID--NTSYLLINCNSTSVITQACP 208
QY 176 ISFEPPIHYCAPAGFALIKNDKTFNGKAPCKNVSTYQCHGRPVVSTOLLNGSLAE 235
DB 209 VSFEPPIHYCA-RWFALINCNNKFNSTGCTVSTVQCHGRPVVSTHLLNGSLAE 267
QY 236 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNGA-----GDRIQ 277
DB 268 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTTSHIGPRAFYATGDIIGDRIQ 327
QY 278 AHCNISRAKNDTLKQIYIKLEQFENKTIVFNHSSGDDPEIVHNSFNGCGEFPYCSTQ 337
DB 328 AHCNISRAKNDTLKQIYIKLEQFENKTIIIFNRSAGDDPEIVHNSFNGCGEFPYCSTQ 387
QY 338 LFNSTNNNTTSGSNNTBEN--TTLPCRKQIINMOEYKAMVAPPIGQIRCSNITGL 396
DB 388 LFNSTN-NGTEGSSNTGNDTITLPCRKEIINMOEYKAMVAPPIGQIRCSNITGL 446
QY 397 LITRDGINENG-----TEIFRPGGDMRDNMRSSELYKYVYKIEPLGVAPTKCKRRVYQ 451
DB 447 LITRDGNSKNKSKNKTETIFRPGGDMRDNMRSSELYKYVYKIEPLGVAPTKCKRRVYQ 506
QY 452 REKRAVG-IGAFLGAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 510
DB 507 REKRAVGITIGAMFLGAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 566
QY 511 QLTWGIKQOLARVLAVERYLGDQQLGIGWCSGKLTCTTAVPNNASMSNKSLEDRINNM 570
DB 567 QLTWGIKQOLARVLAVERYLGDQQLGIGWCSGKLTCTTAVPNNASMSNKSLEDRINNM 626
QY 571 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNPEDITKMLWY 625
DB 627 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNPEDITKMLWY 681

```

RESULT 8  
US-08-127-499A-14  
Sequence 14, Application US/08127499A  
Patent No. 5510264

GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127.499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861 amino acids

TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-14

Query Match 86.6%; Score 2901.5; DB 1; Length 861;  
Best Local Similarity 84.3%; Pred. No. 5.5e-242;  
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

QY 2 EKLWVTVYGVVWKEATITLFCASDAKAYDTEVHNAWMAHACPTDPNPOEVLNANTE 61
DB 32 EKLWVTVYGVVWKEATITLFCASDAKAYDTEVHNAWMAHACPTDPNPOEVLNANTE 91
QY 62 HNNMKNNMTEQWQEDIIISLDQSLKPCVLTPLCVTLNCKDV--NATNT---TNDSEG- 115
DB 92 NFNWKKNDWVEQWHEDIISLDQSLKPCVLTPLCVSLKCTDGLNATNTNSNTSSGE 151
QY 116 -TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
DB 152 MMEKGEIKNCSFNITTSIRDEVOKEYALFYKLDIIPIDNTTSYTLSCNTSVITQACP 211
QY 175 KISFEPPIHYCAPAGFALIKNDKTFNGKAPCKNVSTYQCHGRPVVSTOLLNGSLA 234
DB 212 KISFEPPIHYCAPAGFALIKNDKTFNGCTVSTVQCHGRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
DB 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIIRIGPRAFYATGDIIGDRIQ 331
QY 276 ROAHNISRAKNDTLKQIYIKLEQF-ENKTIVFNHSSGDDPEIVHNSFNGCGEFPYC 334
DB 332 ROAHNISRAKNDTLKQIYIKLEQFENKTIIIFKSSGDDPEIVHNSFNGCGEFPYC 391
QY 335 STOLFNSTFWSTSTGSSNNTBSDTITLPCRKQIINMOEYKAMVAPPIGQIRCS 390
DB 392 STOLFNSTFWSTSTGSSNNTBSDTITLPCRKQIINMOEYKAMVAPPIGQIRCS 451
QY 391 SNITGLLITRDGINENGTEIFRPGGDMRDNMRSSELYKYVYKIEPLGVAPTKCKRRV 450
DB 452 SNITGLLITRDGINENGTEIFRPGGDMRDNMRSSELYKYVYKIEPLGVAPTKCKRRV 511
QY 451 QREKRAVGIGAFLGAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 510
DB 512 QREKRAVGIGAFLGAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 571
QY 511 QLTWGIKQOLARVLAVERYLGDQQLGIGWCSGKLTCTTAVPNNASMSNKSLEDRINNM 570
DB 572 QLTWGIKQOLARVLAVERYLGDQQLGIGWCSGKLTCTTAVPNNASMSNKSLEDRINNM 631
QY 571 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNPEDITKMLWY 625
DB 632 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNPEDITKMLWY 686

```

RESULT 9  
US-08-482-847-14  
Sequence 14, Application US/08482847  
Patent No. 5556757

GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:

```

/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/482,847
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/127,499
/ FILING DATE: 28-SEP-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 51916/104/INBI
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)672-5300
/ TELEFAX: (202)672-5399
/ TELETYPE: 904136
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 861 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: unknown
/ US-08-482-847-14

```

Query Match 86.6%; Score 2901.5; DB 1; Length 861;

Best Local Similarity 84.3%; Pred. No. 5.5e-242; Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

QY 2 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPOEVLNTE 61
DB 32 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPOEVLNTE 91
QY 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 151
QY 116 -TMRGEIKNGSFNTTTSIRDEVQKVALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 174
DB 152 MMEKGEIKNGSFNTTTSIRDEVQKVALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 211
QY 175 KISFEPPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTHGIRPVVSTOLLNGSLA 234
DB 212 KVSFEPPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTTIIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 272 EEEVIRSDNFTNNAKTTIIIVOLKESVEINCTRPNNN-----GAGDI 331
QY 276 ROAHNINISAKKNATLKOIVILRQPF-BNKTIIVRHSGSGPELYMHSFNGGEPFYCN 334
DB 332 ROAHNINISAKKNATLKOIVILRQPF-BNKTIIVRHSGSGPELYMHSFNGGEPFYCN 391
QY 335 STOLFNSTWNN---NTEGSGNTEG-NTITLPCRIKQIIMMOBVGKAMVAPPRIQIRCS 390
DB 392 STOLFNSTWNN---NTEGSGNTEG-NTITLPCRIKQIIMMOBVGKAMVAPPRIQIRCS 451
QY 391 SNITGLLLTRDGGINENGTEIFRPGGDMRDNRSELVYKVKYKLEPLGVAFTKCRVY 450
DB 452 SNITGLLLTRDGGINENGTEIFRPGGDMRDNRSELVYKVKYKLEPLGVAFTKCRVY 511
QY 451 OREKAVGIGAFGLFLAAGSTMGAASTLTVOARLLLSGIVOOONMLLRIEIQOHL 510
DB 512 OREKAVGIGAFGLFLAAGSTMGAASTLTVOARLLLSGIVOOONMLLRIEIQOHL 571
QY 511 QLTWVGICQLOARVLAVERVYLDQOLGIGWSSGKLICTTAVPMNASNGSKSLDRITMNM 570
DB 572 QLTWVGICQLOARVLAVERVYLDQOLGIGWSSGKLICTTAVPMNASNGSKSLDRITMNM 631
QY 571 TMEWEIREIDNTSEIYTLIESONQOEKNEQELLELDKMASLNMWFDITKWLWY 625
DB 625 TMEWEIREIDNTSEIYTLIESONQOEKNEQELLELDKMASLNMWFDITKWLWY 686

```

DB 632 TMEWEIREIDNTSEIYTLIESONQOEKNEQELLELDKMASLNMWFDITKWLWY 686

```

RESULT 10
US-07-956-483-10
/ Sequence 10, Application US/07956483
/ Patent No. 6261799
/ GENERAL INFORMATION:
/ APPLICANT: KIENY, Marie-Paule
/ TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
/ NUMBER OF SEQUENCES: 26
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burns, Doane, Swecker & Mathis
/ STREET: P. O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/956,483
/ FILING DATE: 31-DEC-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO 92/19742
/ FILING DATE: 12-NOV-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 91 05392
/ FILING DATE: 02-MAY-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crane-Feuzy, Sharon B
/ REGISTRATION NUMBER: 36,113
/ REFERENCE/DOCKET NUMBER: 017753-005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 861 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-07-956-483-10

```

Query Match 86.6%; Score 2901.5; DB 3; Length 861;

Best Local Similarity 84.3%; Pred. No. 5.5e-242; Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

QY 2 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPOEVLNTE 61
DB 32 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPOEVLNTE 91
QY 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 151
QY 116 -TMRGEIKNGSFNTTTSIRDEVQKVALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 174
DB 152 MMEKGEIKNGSFNTTTSIRDEVQKVALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 211
QY 175 KISFEPPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTHGIRPVVSTOLLNGSLA 234
DB 212 KVSFEPPIHYCAPAGFALIKCNDKTFNGKXPCKNVSTXQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTTIIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 272 EEEVIRSDNFTNNAKTTIIIVOLKESVEINCTRPNNN-----GAGDI 331

```

```

QY 276 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 334
D 332 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 391
QY 335 STOLFNSSTWNN--NTEGSSNTEG-NTITLPCRIKQIIMMOVBGKAMVAPPIRGQIRCS 390
D 392 STOLFNSSTWNN--NTEGSSNTEG-NTITLPCRIKQIIMMOVBGKAMVAPPIRGQIRCS 451
QY 391 SNITGLLLTDDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTKCRRV 450
D 452 SNITGLLLTDDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTKCRRV 511
QY 451 QREKRAVIGAVLFGFGAAGSTMGASMTLTVQARLLSGIYOQNNLLRAIEAOQRL 510
D 512 QREKRAVIGAVLFGFGAAGSTMGASMTLTVQARLLSGIYOQNNLLRAIEAOQRL 571
QY 511 QLTWVGIKQLOARILAVERYLKDQQLIGWCGSGKILCTTAVPWNASWSKSLDRINNM 570
D 572 QLTWVGIKQLOARILAVERYLKDQQLIGWCGSGKILCTTAVPWNASWSKSLDRINNM 631
QY 571 TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 625
D 632 TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 686

```

## RESULT 11

```

US-08-472-240A-1
: Sequence 1, Application US/08472240A
: Patent No. 6284248
: GENERAL INFORMATION:
: APPLICANT: KIENY, Marie-Paule
: TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/472,240A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/956,483
: FILING DATE: 31-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Teskin, Robin L.
: REGISTRATION NUMBER: 35,030
: REFERENCE/DOCKET NUMBER: 017753-055
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-6620
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 661 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: Proteiin
: LOCATION: 1..854
: US-08-472-240A-1

```

Query Match 86.6%; Score 2901.5; DB 3; Length 861;  
 Best Local Similarity 84.3%; Pred. No. 5,5e-242;  
 Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

QY 2 EKLMTVYGVVYVWKEATTTLPQASDAKAYDTEVHNVATHACVPTDPNPOEVLENVTE 61
D 32 EKLMTVYGVVYVWKEATTTLPQASDAKAYDTEVHNVATHACVPTDPNPOEVLENVTE 91
QY 62 HFMWKNWVEQOEEDISLMDQSLKPCVLTPLCYLNCQV--NATNT---TNDSEG- 115
D 92 HFMWKNWVEQOEEDISLMDQSLKPCVLTPLCYLNCQVLTPLCYLNCQVLTPLCYLNCQV 151
QY 116 -TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 174
D 152 TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 211
QY 175 KISPEPIPIHYCAPAGAILKNDKTPNGKPCKNSTXCTGIRPVSTOLLNGSLA 234
D 212 KISPEPIPIHYCAPAGAILKNDKTPNGKPCKNSTXCTGIRPVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNN--GAGDI 275
D 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNN--GAGDI 331
QY 276 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 334
D 332 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 391
QY 335 STOLFNSSTWNN--NTEGSSNTEG-NTITLPCRIKQIIMMOVBGKAMVAPPIRGQIRCS 390
D 392 STOLFNSSTWNN--NTEGSSNTEG-NTITLPCRIKQIIMMOVBGKAMVAPPIRGQIRCS 451
QY 391 SNITGLLLTDDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTKCRRV 450
D 452 SNITGLLLTDDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTKCRRV 511
QY 451 QREKRAVIGAVLFGFGAAGSTMGASMTLTVQARLLSGIYOQNNLLRAIEAOQRL 510
D 512 QREKRAVIGAVLFGFGAAGSTMGASMTLTVQARLLSGIYOQNNLLRAIEAOQRL 571
QY 511 QLTWVGIKQLOARILAVERYLKDQQLIGWCGSGKILCTTAVPWNASWSKSLDRINNM 570
D 572 QLTWVGIKQLOARILAVERYLKDQQLIGWCGSGKILCTTAVPWNASWSKSLDRINNM 631
QY 571 TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 625
D 632 TWMEEREINDYSEIYTLIEESONOEKNEQELLEDKMASLNMNFDITKMLY 686

```

## RESULT 12

```

US-08-472-240A-7
: Sequence 7, Application US/08472240A
: Patent No. 6284248
: GENERAL INFORMATION:
: APPLICANT: KIENY, Marie-Paule
: TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/472,240A
: FILING DATE: 07-JUN-1995

```



```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,483
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..831
DB-472-240A-7

```

```

Query Match      86.6%; Score 2901.5; DB 3; Length 861;
Best Local Similarity 84.3%; Pred. No. 5.5e-242;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

```

QY 2 EKLAWTVVYGVVWKEATTTLCASDAKAYDTEVHNWATACVPTDNPQEVLENTE 61
DB 32 EKLAWTVVYGVVWKEATTTLCASDAKAYDTEVHNWATACVPTDNPQEVLENTE 91
QY 62 HNNMKNMNVQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NFNMMKNDVQWQEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSSGE 151
QY 116 -TMEGEIKNSFNITTSIRDEVQKAYLFYKLDVXVXIDNNNTSYRLISCDTSVITQACP 174
DB 152 MMEKEGEEKNSFNISTIRGVQKAYAFYKLDIIPINDNTSYRLISCDTSVITQACP 211
QY 175 KISFPIPIHYCAPGAFALCKNDKTFNGKXPKKVVSTXQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFPIPIHYCAPGAFALCKNNKTFNGTGPCTVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVVISRDNFTNNAKTIIVQLKESYEINCTRPNNN-----GAGDI 275
DB 272 EEEVVISRDNFTNNAKTIIVQLKESYEINCTRPNNNTRKIRIQRGPRGAFVITIGKGM 331
QY 276 ROAHGNISRAKKNATLKQIVIKLRQF-ENKTIIVNHSNGDPEIVMHSFNGGSEFFYCN 334
DB 332 ROAHGNISRAKKNATLKQIVIKLRQF-ENKTIIVNHSNGDPEIVMHSFNGGSEFFYCN 391
QY 335 STQLFNSTWNN---NTEGSNNTEG-NTTLPRIQIINMGOEVGRKAWAPIRIQIICS 390
DB 392 STQLFNSTWNNSTWSTEGSNNTGSDTITLPRIQIINMGOEVGRKAWAPIRIQIICS 451
QY 391 SNITGLILTRDGIENGTETFRPGGDMRDNRSELVYKVKYLEPLGVAATYCKRRV 450
DB 452 SNITGLILTRDGIENGTETFRPGGDMRDNRSELVYKVKYLEPLGVAATYCKRRV 511
QY 451 OREKRAVIGAVLGLGAGSTWGAASWTLTVQARLLLSGIVQOONNLLRAIEAQOHL 510
DB 512 OREKRAVIGAVLGLGAGSTWGAASWTLTVQARLLLSGIVQOONNLLRAIEAQOHL 571
QY 511 QLTWVGIGLOLQARVLAVERLGDQOLLGWGSGGLICTAVPNNASNSKSLDIAIWM 570
DB 572 QLTWVGIGLOLQARVLAVERLGDQOLLGWGSGGLICTAVPNNASNSKSLDIAIWM 631
QY 571 TMEWEERIDNVTSEIYTLIESQNOQEKNEOELLEDKASIMWPFDTTKLWY 625
DB 632 TMEWEERIDNVTSEIYTLIESQNOQEKNEOELLEDKASIMWPFDTTKLWY 686

```

RESULT 13

```

US-08-817-441-103
Sequence 103, Application US/08817441
Patent No. 6399294
GENERAL INFORMATION:
APPLICANT: CHARBEAU, PIERRE
APPLICANT: CLAVEL, FRANCOISE
APPLICANT: BORMAN, ANDREW
APPLICANT: OUILLENT, CAROLINE
APPLICANT: GUETARD, DENISE
APPLICANT: MONTAGNIER, LUC
APPLICANT: DONJON DE SAINT-MARTIN, JOCELINE
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
NUMBER OF INVENTION: SUBTYPE) ANTIGENS
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSER: Finnegan, Henderson, Farabow, Garrett &
ADDRESSER: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,441
FILING DATE: 11-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 95/01391
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: FR 9412554
FILING DATE: 20-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9502526
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03260, 6005-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-441-103

```

```

Query Match      86.6%; Score 2901.5; DB 4; Length 861;
Best Local Similarity 84.3%; Pred. No. 5.5e-242;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

```

QY 2 EKLAWTVVYGVVWKEATTTLCASDAKAYDTEVHNWATACVPTDNPQEVLENTE 61
DB 32 EKLAWTVVYGVVWKEATTTLCASDAKAYDTEVHNWATACVPTDNPQEVLENTE 91
QY 62 HNNMKNMNVQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NFNMMKNDVQWQEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSSGE 151
QY 116 -TMEGEIKNSFNITTSIRDEVQKAYLFYKLDVXVXIDNNNTSYRLISCDTSVITQACP 174
DB 152 MMEKEGEEKNSFNISTIRGVQKAYAFYKLDIIPINDNTSYRLISCDTSVITQACP 211
QY 175 KISFPIPIHYCAPGAFALCKNDKTFNGKXPKKVVSTXQCTHGIRPVVSTOLLNGSLA 234

```

```

Db 212 KVSFPIPIHYCAPAGFAILKKNNTFNGTGPCTVSTVQCTHGIRPVSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
Db 272 EEEVIRSANFTDAKTIIVOLNOSVEINCTRPNNNRKSIIRIORGPRAVTIGKIGNM 331
Qy 276 ROAHNISRAKNDTLKOIYIKLEOF-ENKTIIVFNHSSGDDPEIVMHSFNGGGEFFYCN 334
Db 332 ROAHNISRAKNDTLKOIYIKLEOF-ENKTIIVFNHSSGDDPEIVMHSFNGGGEFFYCN 391
Qy 335 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 390
Db 392 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 451
Qy 391 SNITGLLTRDGINENGTETFRPGGDMRDNRSELYKRVKIEPLGVAPTAKRRV 450
Db 452 SNITGLLTRDGINENGTETFRPGGDMRDNRSELYKRVKIEPLGVAPTAKRRV 511
Qy 451 OREKRAVGIGAVFLGFLAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRM 510
Db 512 OREKRAVGIGAVFLGFLAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRM 571
Qy 511 QLTWGIKQOLARILAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSLEDRIMNM 570
Db 572 QLTWGIKQOLARILAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSLEDRIMNM 631
Qy 571 TWMEEREIDNVTSEIYTLIESONQOEKNEBELLEDKNASLNMWFDITKLMY 625
Db 632 TWMEEREIDNVTSEIYTLIESONQOEKNEBELLEDKNASLNMWFDITKLMY 686

```

## RESULT 14

```

US-07-956-483-16
; Sequence 16, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: 9P160 VARIANT
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-956-483-16

```

```

Query Match 86.3%; Score 2893.5; DB 3; Length 861;
Best Local Similarity 84.1%; Pred. No. 2,7e-241;
Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

```

Qy 2 EKLWVTVYGVPPWKEATITLFCASDAKAYDTEHANNMATHACVPTDPNQEVLNVTE 61
Db 32 EKLWVTVYGVPPWKEATITLFCASDAKAYDTEHANNMATHACVPTDPNQEVLNVTE 91
Qy 62 HNNMKNMVEQOMEDIISLMDOSLKPCVLTPLCVTLNCKDV-NATNT---TNDSEG- 115
Db 92 HNNMKNMVEQOMEDIISLMDOSLKPCVLTPLCVTLNCKDV-NATNT---TNDSEG- 151
Qy 116 -TWGEIKKCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISGDTSVITQAC 174
Db 152 MMMEKGEIKKCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISGDTSVITQAC 211
Qy 175 KISFPIPIHYCAPAGFAILKKNNTFNGTGPCTVSTVQCTHGIRPVSTOLLNGSLA 234
Db 212 KVSFPIPIHYCAPAGFAILKKNNTFNGTGPCTVSTVQCTHGIRPVSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
Db 272 EEEVIRSANFTDAKTIIVOLNOSVEINCTRPNNNRKSIIRIORGPRAVTIGKIGNM 331
Qy 276 ROAHNISRAKNDTLKOIYIKLEOF-ENKTIIVFNHSSGDDPEIVMHSFNGGGEFFYCN 334
Db 332 ROAHNISRAKNDTLKOIYIKLEOF-ENKTIIVFNHSSGDDPEIVMHSFNGGGEFFYCN 391
Qy 335 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 390
Db 392 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 451
Qy 391 SNITGLLTRDGINENGTETFRPGGDMRDNRSELYKRVKIEPLGVAPTAKRRV 450
Db 452 SNITGLLTRDGINENGTETFRPGGDMRDNRSELYKRVKIEPLGVAPTAKRRV 511
Qy 451 OREKRAVGIGAVFLGFLAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRM 510
Db 512 OREKRAVGIGAVFLGFLAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRM 571
Qy 511 QLTWGIKQOLARILAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSLEDRIMNM 570
Db 572 QLTWGIKQOLARILAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSLEDRIMNM 631
Qy 571 TWMEEREIDNVTSEIYTLIESONQOEKNEBELLEDKNASLNMWFDITKLMY 625
Db 632 TWMEEREIDNVTSEIYTLIESONQOEKNEBELLEDKNASLNMWFDITKLMY 686

```

## RESULT 15

```

US-09-124-900-9
; Sequence 9, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PUTSCHER, Martin
; APPLICANT: TROLOA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHWARTZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P

```

```

1 CURRENT APPLICATION NUMBER: US/09/124,900
2 CURRENT FILING DATE: 1998-07-30
3 PRIOR APPLICATION NUMBER: PCT/EP95/01481
4 PRIOR FILING DATE: 1995-04-19
5 NUMBER OF SEQ ID NOS: 11
6 SOFTWARE: PatentIn version 3.0
7 SEQ ID NO: 9
8 LENGTH: 856
9 TYPE: PR1
10 ORGANISM: Human immunodeficiency virus type 1
11 FEATURE:
12 NAME/KEY: PEPTIDE
13 LOCATION: (79)..(184)
14 NAME/KEY: PEPTIDE
15 LOCATION: (326)..(400)
16
17 US-09-124-900-9

```

Query Match	86.2%;	Score 2887;	DB 3;	length 856;
Best Local Similarity	Pred. 84.0%;	No. 9.7e-241;		
Matches 546;	Conservative 34;	Mismatches 44;	Indels 26;	Gaps 5

Db	32	EKLMTVYVGVPMWKATATTLPCASAPAKAYDVEHNWATACVPTDNPBEVLAVNTE	61
Qy	62	HFNMKNMVEQOEODIISLMDOSLKPCYKLTPLCVTLNCKDVNATNTDSEGTMER	119
Db	92	NFMWKNMDWEOQHEDIIISLMDOSLKPCYKLTPLCVSLKCTDLDKNDTNTSSGRMIMTEK	151
Qy	120	GEIKGCSFNITTSIRDEVOKEVALPKJLWVXIDNNNTSYRLISDTSVITQACPXSFE	179
Db	152	GEIKGCSFNISTISRGVQKEVAFPKJLDIPIDDITTSYLTSCNSVITQACPXSFE	211
Qy	180	PIPIHYCAPAGFALILCKNDKTFNGKXPCXNVSTOCTHGIRPVYSTOLLNGSLAEBEV	239
Db	212	PIPIHYCAPAGFALILCKNKTENGSTGPTCNVSTVQCTHGIRPVYSTOLLNGSLAEBEV	271
Qy	240	IRSDNFTNNAKTIIVOLKESVENCIRPNNN-----GAGDIRQAH	286
Db	272	IRANLFDNAKTIIVOLNOSVENCIRPNNNTRKSIRIORGBRAFTVIGIKIMMRQAH	331
Qy	281	NISRAKANDLKOIVILKREOF-ENTTYIYFNHSSGDPPEIYVHSPNCGEFPYCNSTOLF	339
Db	332	NISRAKANNLKOIDSKLREQFOGNKTIIFKXSSGDEPEIYTHSPNCGEFPYCNSTOLF	391
Qy	340	NSTMWN---NTEGSNNTEG-NTITLPCRIKOIINMVOEVKAMVAPIRGOIRCSSNITG	399
Db	392	NSTMFNSTWSTKGSNNTEGSNTITLPCRIKOIINMVOEVKAMVAPIRGOIRCSSNITG	451
Qy	396	LLLTROGGINENGETEIPREGGDMRDNMSELYKTKVYKIEPLGVAPYKCRARYOBEKR	455
Db	452	LLLTROGGSNNSEEIFREGGDMRDNMSELYKTKVYKIEPLGVAPYKARRRVOREKR	511
Qy	456	AVGIGAVFLGFLAAGASTGASMTLTVOARLLSGIYOQONNLRAIEAQOHWLOLTVW	515
Db	512	AVGIGALFLGFLGASSTGASMTLTVOAROLLSGIYOQONNLRAIEAQOHLLOLTVW	571
Qy	516	GIKOLOARVLAVERYLGDQOLLGIWCGSGKLICTTAVPMNWSWSNKSIDRIINMNTWMEW	579
Db	572	GIKOLOARILIAVERYLKDQOLLGIWCGSGKLICTTAVPMNWSWSNKSLEQIINMNTWMEW	631
Qy	576	EREIDNNTSEIYLTIEESONOEKNOBELLEDKASILMNPDTKMWY	625
Db	632	DREINNTSLIHSIEESONOEKNOBELLEDKASILMNPENITNLMWY	681

Search completed: December 12, 2003, 12:34:29  
Job time : 15.2122 secs

**THIS PAGE BLANK (08.70)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 26.0558 Seconds  
(without alignments)  
4461.192 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351  
Sequence: 1 VEKLMVTYYGVVWKREATT.....EIDKMASIMNPFITKMLMY 625

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	99.8	625	US-10-032-162-17	Sequence 17, Appli
2	3295	98.3	643	US-10-032-162-13	Sequence 13, Appli
3	2928	87.4	842	US-10-190-435-2	Sequence 2, Appli
4	2928	87.4	842	US-10-241-009-2	Sequence 2, Appli
5	2928	87.4	842	US-10-190-434B-2	Sequence 2, Appli
6	2928	87.4	842	US-10-190-305A-2	Sequence 2, Appli
7	2928	87.4	847	US-09-476-242-2	Sequence 2, Appli
8	2923	87.4	579	US-10-032-162-15	Sequence 15, Appli
9	2901.5	86.6	861	US-10-026-741-103	Sequence 103, App
10	2878	85.9	856	US-09-476-242-1	Sequence 1, Appli
11	2867	85.6	856	US-10-196-515-11	Sequence 11, Appli
12	2822.5	84.2	726	US-10-196-515-3	Sequence 3, Appli
13	2815.5	84.0	759	US-10-196-515-12	Sequence 12, Appli
14	2793.5	83.4	868	US-09-938-406-1	Sequence 1, Appli
15	2758	82.3	619	US-09-891-609-4	Sequence 4, Appli

16	2758	82.3	646	US-09-891-609-2	Sequence 2, Appli
17	2611.5	77.9	860	US-10-190-435-6	Sequence 6, Appli
18	2611.5	77.9	860	US-10-241-009-6	Sequence 6, Appli
19	2611.5	77.9	860	US-10-190-434B-6	Sequence 6, Appli
20	2611.5	77.9	860	US-10-190-305A-6	Sequence 6, Appli
21	2572	76.8	867	US-10-190-435-3	Sequence 3, Appli
22	2572	76.8	867	US-10-190-435-126	Sequence 126, App
23	2572	76.8	867	US-10-241-009-3	Sequence 3, Appli
24	2572	76.8	867	US-10-190-434B-3	Sequence 3, Appli
25	2572	76.8	867	US-10-190-305A-3	Sequence 3, Appli
26	2568	76.6	855	US-10-369-294-11	Sequence 11, Appli
27	2568	76.6	855	US-10-369-294-12	Sequence 12, Appli
28	2565.5	76.6	803	US-10-190-435-135	Sequence 135, App
29	2564.5	76.5	803	US-10-190-435-134	Sequence 134, App
30	2563	76.5	869	US-10-190-435-4	Sequence 4, Appli
31	2563	76.5	869	US-10-241-009-4	Sequence 4, Appli
32	2563	76.5	869	US-10-190-434B-4	Sequence 4, Appli
33	2563	76.5	869	US-10-190-305A-4	Sequence 4, Appli
34	2562	76.5	853	US-10-369-294-13	Sequence 13, Appli
35	2562	76.5	861	US-10-190-435-139	Sequence 139, App
36	2560	76.4	858	US-10-190-435-150	Sequence 150, App
37	2559.5	76.4	870	US-10-190-435-127	Sequence 127, App
38	2553	76.2	870	US-10-190-435-147	Sequence 147, App
39	2552	76.2	862	US-10-190-435-141	Sequence 141, App
40	2551.5	76.1	865	US-10-190-435-140	Sequence 140, App
41	2550.5	76.1	845	US-10-190-435-129	Sequence 129, App
42	2550.5	76.1	845	US-10-190-435-130	Sequence 130, App
43	2546.5	76.0	857	US-10-190-435-138	Sequence 138, App
44	2542.5	75.9	855	US-10-190-435-144	Sequence 144, App
45	2537.5	75.7	859	US-10-190-435-137	Sequence 137, App

#### ALIGNMENTS

RESULT 1  
US-10-032-162-17  
; Sequence 17, Application US/10032162  
; Publication No. US20030052839A1  
GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331az  
; CURRENT APPLICATION NUMBER: US/10/032.162  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: 09/602,864  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 625  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (151)..(151)  
; OTHER INFORMATION: X-UNKNOWN AMINO ACID  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (205)..(205)  
; OTHER INFORMATION: X-UNKNOWN AMINO ACID  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (213)..(213)  
; OTHER INFORMATION: X-UNKNOWN AMINO ACID  
US-10-032-162-17  
Query Match 99.8%; Score 3345; DB 15; Length 625;  
Best Local Similarity 100.0%; Pred. No. 9.3e-311;

Matches	617;	Conservative	0;	Mismatches	8;	Indels	18;	Gaps	1
Qy	1	VEKLVMTVYVYGVPVWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVL	ENVVT	60					
Db	1	VEKLVMTVYVYGVPVWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVL	ENVVT	60					
Qy	61	EHEFMWKNWNEQOEEDILSLMDSIKRVCVLTPLCVTLNCKDNVA	TNTNDS	ECTMERG	120				
Db	61	EHEFMWKNWNEQOEEDILSLMDSIKRVCVLTPLCVTLNCKDNVA	TNTNDS	ECTMERG	120				
Qy	121	EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDT	SYITQAC	PKRISFEP	180				
Db	121	EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDT	SYITQAC	PKRISFEP	180				
Qy	181	IPHYCAPAGFALLCKNDKTENGKXPCKNVSTXOCTHGIRPVVSTOLL	NGSLAE	EEVVI	240				
Db	181	IPHYCAPAGFALLCKNDKTENGKXPCKNVSTXOCTHGIRPVVSTOLL	NGSLAE	EEVVI	240				
Qy	241	RSNDFTNNAKTIIVQLKESVEINCTRPNNG-----AGDIRAHQNI		282					
Db	241	RSNDFTNNAKTIIVQLKESVEINCTRPNNRKSIIHGPRAFYTTGELIG	DIRAHQNI	300					
Qy	283	SRAKNDTLKOIVIKLREOFENKTIIVFNHSSGSDPEIVMHSFNC	GGEFFY	CNSTOLL	FNST	342			
Db	301	SRAKNDTLKOIVIKLRQFENKTIIVFNHSSGSDPEIVMHSFNC	GGEFFY	CNSTOLL	FNST	360			
Qy	343	WNNTEGSNNTEGNTTLPCKRIKQIINNMOEYKAMVAPPIRGQIR	CSNITGL	LLTRDG	402				
Db	361	WNNTEGSNNTEGNTTLPCKRIKQIINNMOEYKAMVAPPIRGQIR	CSNITGL	LLTRDG	420				
Qy	403	GINNGNTEIFPPGGDMDNRSELYKKVKKIETPLGVAPTKCKR	VVORE	BAVGIGAV	462				
Db	421	GINNGNTEIFPPGGDMDNRSELYKKVKKIETPLGVAPTKCKR	VVORE	BAVGIGAV	480				
Qy	463	FLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNNLRAIE	QOQRMQL	FWYGIKQLOA	522				
Db	481	FLGFLGAAGSTMGAAASMTLTVOARLLLSGIVQOQNNLRAIE	QOQRMQL	FWYGIKQLOA	540				
Qy	523	RLVAVERYLGDQOLLGIWCGSGKLICTTAVPMNASWSKSLDR	IWNNT	WMEWERIDNY	582				
Db	541	RLVAVERYLGDQOLLGIWCGSGKLICTTAVPMNASWSKSLDR	IWNNT	WMEWERIDNY	600				
Qy	583	TSEIYTLIEESQNOQEKNEOELLEIDKASLIMNWDTIKWLMY		625					
Db	601	TSEIYTLIEESQNOQEKNEOELLEIDKASLIMNWDTIKWLMY		643					
RESULT 3									
US-10-190-435-2									
Sequence 2, Application US/10190435									
Publication No. US20030143248A1									
GENERAL INFORMATION:									
APPLICANT: ZUR MEGEDF, Jan									
APPLICANT: BARNETT, Susan W.									
APPLICANT: LIAN, Ying									
APPLICANT: ENGELBRECHT, Susan									
APPLICANT: VAN RENSBURG, Estrelita J.									
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C									
FILE REFERENCE: PPI8133.003 / 2302-18133									
CURRENT APPLICATION NUMBER: US/10/190,435									
CURRENT FILING DATE: 2002-12-30									
NUMBER OF SEQ ID NOS: 319									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 2									
LENGTH: 842									
TYPE: PRT									
ORGAN									

Best Local Similarity 86.0%; Pred. No. 1.1e-270;  
Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

QY 1 VEKLVWTVYGVVPMWEATTTLFCASDAKAYDTEVHNWATACVPTDNPDEVLIENT 60
DB 25 VEKLVWTVYGVVPMWEATTTLFCASDAKAYDTEVHNWATACVPTDNPDEVLIENT 84
QY 61 EHFNNKNNMNVBOMEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 119
DB 85 EHFNNKNNMNVBOMEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 144
QY 120 GEIKNCSFNITTSIRDEVOKAYALFYKLDVYXIDNNNTSYRLISCDTSYITTOACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNKQKEYALFYKLDVYPIDNDNTSYKLINCNTSYITTOACPKVSPSE 204
QY 160 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEVY 239
DB 205 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEVY 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHGN 281
DB 265 IRSENFPTNDAKTIIVOLKESVEINCTRPNNNTNRKSIITIGPRAFYATGDIIGDIRQAHGN 324
QY 282 ISRAKMDTLKQIVIKLBQFENKTIIVFNHSSGGDEIYVHSHFNCGGEFFYCNSTOLFNS 341
DB 345 ISGEKNNNTLKOIVTKLQAFGNKTIIVFKOSSGDEIYVHSHFNCGGEFFYCNSTOLFNS 384
QY 342 TWNNNTBESNTEGNTITLPCRIKQIIMMOWGVGKAMVAPPIRGDIRCSNITGLLTLD 401
DB 365 TW-NNTIGPNNNTNG--TITLPCRIKQIINRMQEVGKAMVAPPIRGDIRCSNITGLLTLD 442
QY 402 GGIN-ENGTEIFRPGCGDMRDNRSELKYKVKVKEIPLGVAATPKCKRVRVQREKAVGIG 460
DB 443 GGEKESINTTEIFRPGCGDMRDNRSELKYKVKVKEIPLGVAATPKCKRVRVQREKAVGIG 502
QY 461 AVFLCPLGAGSTWGAASNTLTVQARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 520
DB 503 AMFLGFLGAGSTWGAASNTLTVQARQLLSGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 562
QY 521 QARVLAVERYLGDQOLLGIGWGSGLICTTAAPNMAASNSKSLDRIINNMTMWEWEREID 580
DB 553 QARVLAVERYLKDQOLLGIGWGSGLICTTAAPNMAASNSKSLDRIINNMTMWEWEREID 622
QY 581 NYTSEIYTLIESONQOEKNEOELBELDKMASLMMNPDTTKWLMY 625
DB 623 NYTNLIYTLIESONQOEKNEOELBELDKMASLMMNPDTTKWLMY 667

```

RESULT 4  
 US-10-241-009-2  
 ; Sequence 2, Application US/10241009  
 ; Publication No. US20030170614A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZUR MEGEDE, Jan  
 ; APPLICANT: BARNETT, Susan  
 ; APPLICANT: LIAN, Ying  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B  
 ; FILE REFERENCE: 2300-1621.21  
 ; CURRENT APPLICATION NUMBER: US/10/241,009  
 ; CURRENT FILING DATE: 2002-12-13  
 ; NUMBER OF SEQ ID NOS: 68  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 842  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SP162  
 US-10-241-009-2

Query Match 87.4%; Score 2928; DB 12; Length 842;  
 Best Local Similarity 86.0%; Pred. No. 1.1e-270;  
 Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

QY 1 VEKLVWTVYGVVPMWEATTTLFCASDAKAYDTEVHNWATACVPTDNPDEVLIENT 60
DB 25 VEKLVWTVYGVVPMWEATTTLFCASDAKAYDTEVHNWATACVPTDNPDEVLIENT 84
QY 61 EHFNNKNNMNVBOMEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 119
DB 85 EHFNNKNNMNVBOMEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 144
QY 120 GEIKNCSFNITTSIRDEVOKAYALFYKLDVYXIDNNNTSYRLISCDTSYITTOACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNKQKEYALFYKLDVYPIDNDNTSYKLINCNTSYITTOACPKVSPSE 204
QY 160 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEVY 239
DB 205 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEVY 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHGN 281
DB 265 IRSENFPTNDAKTIIVOLKESVEINCTRPNNNTNRKSIITIGPRAFYATGDIIGDIRQAHGN 324
QY 282 ISRAKMDTLKQIVIKLBQFENKTIIVFNHSSGGDEIYVHSHFNCGGEFFYCNSTOLFNS 341
DB 345 ISGEKNNNTLKOIVTKLQAFGNKTIIVFKOSSGDEIYVHSHFNCGGEFFYCNSTOLFNS 384
QY 342 TWNNNTBESNTEGNTITLPCRIKQIIMMOWGVGKAMVAPPIRGDIRCSNITGLLTLD 401
DB 365 TW-NNTIGPNNNTNG--TITLPCRIKQIINRMQEVGKAMVAPPIRGDIRCSNITGLLTLD 442
QY 402 GGIN-ENGTEIFRPGCGDMRDNRSELKYKVKVKEIPLGVAATPKCKRVRVQREKAVGIG 460
DB 443 GGEKESINTTEIFRPGCGDMRDNRSELKYKVKVKEIPLGVAATPKCKRVRVQREKAVGIG 502
QY 461 AVFLCPLGAGSTWGAASNTLTVQARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 520
DB 503 AMFLGFLGAGSTWGAASNTLTVQARQLLSGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 562
QY 521 QARVLAVERYLGDQOLLGIGWGSGLICTTAAPNMAASNSKSLDRIINNMTMWEWEREID 580
DB 553 QARVLAVERYLKDQOLLGIGWGSGLICTTAAPNMAASNSKSLDRIINNMTMWEWEREID 622
QY 581 NYTSEIYTLIESONQOEKNEOELBELDKMASLMMNPDTTKWLMY 625
DB 623 NYTNLIYTLIESONQOEKNEOELBELDKMASLMMNPDTTKWLMY 667

```

RESULT 5  
 US-10-190-434B-2  
 ; Sequence 2, Application US/10190434B  
 ; Publication No. US20030194800A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZUR MEGEDE, Jan  
 ; APPLICANT: BARNETT, Susan  
 ; APPLICANT: LIAN, Ying  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B  
 ; FILE REFERENCE: 2300-1621.20  
 ; CURRENT APPLICATION NUMBER: US/10/190,434B  
 ; CURRENT FILING DATE: 2002-07-05  
 ; NUMBER OF SEQ ID NOS: 68  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 842  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SP162  
 US-10-190-434B-2

Query Match 87.4%; Score 2928; DB 12; Length 842;  
 Best Local Similarity 86.0%; Pred. No. 1.1e-270;  
 Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

Qy 1 VEKLMVTVYGVGVWPKKATTTTLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEVLLENVT 60
Db 25 VEKLMVTVYGVGVWPKKATTTTLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEVLLENVT 84
Qy 61 EHFNMKNWVQWQEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 119
Db 85 ENFMWKNWVQWQEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 144
Qy 120 GEINCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPXSFE 179
Db 145 GEINCSFKVTTISRNKQKEVALFYKLDVPIIDNDTSYKILNCTNSVITQACPXSFE 204
Qy 180 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 239
Db 205 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 264
Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 281
Db 265 IRSNFTNNAKTIIVOLKESVEINCTRPNNNRKSIITIGPRAFYATGDIIGDIRQAHN 324
Qy 282 ISRAKNDTLKQIYIKLREOPENKTIVNHSNGGDEIIVMHSFNGGGEFFYCNSTOLFNS 341
Db 325 ISGERMNTLQIYIKLQAFGNKTIIVKQSSGDEIIVMHSFNGGGEFFYCNSTOLFNS 384
Qy 342 TWNNTEGSNTEGNTITLPCRKOIIMMOEVEKAMYPAPIRGOIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNNTNG-TITLPCRKOIINRMOEVEKAMYPAPIRGOIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKVKVIEPLGVAPTKCRKRVVOREKAVGIG 460
Db 443 GKKEISNTTEIFRPGGDMRDNRSELYKVKVIEPLGVAPTKCRKRVVOREKAVGIG 502
Qy 461 AVPLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQUTWGIKOL 520
Db 503 AMFLGFLGAGSTMGASMTLTVQARQLLSGIVQOONNLRAIEAOQMLQUTWGIKOL 562
Qy 521 QARVLAVERYLGDQQLIGWCGSKLICCTAVPNNASMSNKSIDRINNNMTMWEREID 580
Db 563 QARVLAVERYLKDQQLIGWCGSKLICCTAVPNNASMSNKSIDRINNNMTMWEREID 622
Qy 581 NYTSEIYTLIEBSONQOEKNEQELLEDKMASLNNWPDITKWLWY 625
Db 623 NYTMLIYTLIEBSONQOEKNEQELLEDKMASLNNWPDISKWLWY 667

```

## RESULT 6

```

US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-305A-2

```

Query Match 87.4%; Score 2928; DB 12; Length 842;  
 Best Local Similarity 86.0%; Pred. No. 1,1e-270;  
 Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

Qy 1 VEKLMVTVYGVGVWPKKATTTTLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEVLLENVT 60
Db 25 VEKLMVTVYGVGVWPKKATTTTLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEVLLENVT 84
Qy 61 EHFNMKNWVQWQEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 119
Db 85 ENFMWKNWVQWQEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 144
Qy 120 GEINCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPXSFE 179
Db 145 GEINCSFKVTTISRNKQKEVALFYKLDVPIIDNDTSYKILNCTNSVITQACPXSFE 204
Qy 180 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 239
Db 205 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 264
Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 281
Db 265 IRSNFTNNAKTIIVOLKESVEINCTRPNNNRKSIITIGPRAFYATGDIIGDIRQAHN 324
Qy 282 ISRAKNDTLKQIYIKLREOPENKTIVNHSNGGDEIIVMHSFNGGGEFFYCNSTOLFNS 341
Db 325 ISGERMNTLQIYIKLQAFGNKTIIVKQSSGDEIIVMHSFNGGGEFFYCNSTOLFNS 384
Qy 342 TWNNTEGSNTEGNTITLPCRKOIIMMOEVEKAMYPAPIRGOIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNNTNG-TITLPCRKOIINRMOEVEKAMYPAPIRGOIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKVKVIEPLGVAPTKCRKRVVOREKAVGIG 460
Db 443 GKKEISNTTEIFRPGGDMRDNRSELYKVKVIEPLGVAPTKCRKRVVOREKAVGIG 502
Qy 461 AVPLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQUTWGIKOL 520
Db 503 AMFLGFLGAGSTMGASMTLTVQARQLLSGIVQOONNLRAIEAOQMLQUTWGIKOL 562
Qy 521 QARVLAVERYLGDQQLIGWCGSKLICCTAVPNNASMSNKSIDRINNNMTMWEREID 580
Db 563 QARVLAVERYLKDQQLIGWCGSKLICCTAVPNNASMSNKSIDRINNNMTMWEREID 622
Qy 581 NYTSEIYTLIEBSONQOEKNEQELLEDKMASLNNWPDITKWLWY 625
Db 623 NYTMLIYTLIEBSONQOEKNEQELLEDKMASLNNWPDISKWLWY 667

```

## RESULT 7

```

US-09-476-242-2
; Sequence 2, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-2

```

Query Match 87.4%; Score 2928; DB 10; Length 847;  
 Best Local Similarity 86.0%; Pred. No. 1,1e-270;  
 Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;



QY 61 EHFNNKNNNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMER 119  
 DB 90 EHFNNKNNNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMER 149  
 QY 120 GGIKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFE 179  
 DB 150 GGIKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFE 209  
 QY 180 PPIHYCAPAGFALLKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVVI 239  
 DB 210 PPIHYCAPAGFALLKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVVI 269  
 QY 240 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCH 281  
 DB 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCH 329  
 QY 282 ISRAKNDTLKOIVILKREOFENKTIIVFNHSSGDEBEIVMHSFNCGEFFYCNSDTLPNST 341  
 DB 330 ISRAKNDTLKOIVILKREOFENKTIIVFNHSSGDEBEIVMHSFNCGEFFYCNSDTLPNST 389  
 QY 342 TNNNTGSSNTEGNTITLPCRIKQIINMMOEVGKAMVAPPIRGQIRCSSNTITGLLTRDG 401  
 DB 390 TNNNTGSSNTEGNTITLPCRIKQIINMMOEVGKAMVAPPIRGQIRCSSNTITGLLTRDG 447  
 QY 402 GGIN-ENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 460  
 DB 448 GGIN-ENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 507  
 QY 461 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQRMQLTVMGIGKOLA 520  
 DB 508 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQRMQLTVMGIGKOLA 567  
 QY 521 QARVAVERYLGDQOOLGIGWCSGKLICTTAVPNNASWSKSLDRINNNMTMEMEREID 580  
 DB 568 QARVAVERYLGDQOOLGIGWCSGKLICTTAVPNNASWSKSLDRINNNMTMEMEREID 627  
 QY 581 NYTSEIYTLIEESONQOEKNEOELLELDKVASLMMWFDITKWLWY 625  
 DB 628 NYTSEIYTLIEESONQOEKNEOELLELDKVASLMMWFDITKWLWY 672

RESULT 8  
 US-10-032-162-15  
 ; Sequence 15, Application US/10032162  
 ; Publication No. US20030052839A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BINLEY, JAMES M  
 ; APPLICANT: SCHUELEK, NORBERT  
 ; APPLICANT: OLSON, WILLIAM C  
 ; APPLICANT: PAUL, MADDON J  
 ; APPLICANT: JOHN, MOORE P  
 ; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 2048/59331az  
 ; CURRENT APPLICATION NUMBER: US/10/032.162  
 ; CURRENT FILING DATE: 2002-10-24  
 ; PRIOR APPLICATION NUMBER: 09/602,864  
 ; PRIOR FILING DATE: 2000-06-23  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 15  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Human immunodeficiency virus type 1  
 US-10-032-162-15

Query Match 87.2%; Score 2923; DB 15; Length 579;  
 Best Local Similarity 86.0%; Pred. No. 1.9e-270;  
 Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;  
 QY 1 VEKLVVVVYGVVWKEATTTIFCASDAKAYTEVHNWATACVPTDNPQEVLENT 60  
 DB 1 VEKLVVVVYGVVWKEATTTIFCASDAKAYTEVHNWATACVPTDNPQEVLENT 60

QY 61 EHFNNKNNNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMER 120  
 DB 61 EHFNNKNNNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMER 98  
 QY 121 EIKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFE 180  
 DB 99 -----GCDTSVITQACPKISFE 116  
 QY 181 PPIHYCAPAGFALLKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVVI 240  
 DB 117 PPIHYCAPAGFALLKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVVI 176  
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCH 282  
 DB 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCH 236  
 QY 283 ISRAKNDTLKOIVILKREOFENKTIIVFNHSSGDEBEIVMHSFNCGEFFYCNSDTLPNST 342  
 DB 237 ISRAKNDTLKOIVILKREOFENKTIIVFNHSSGDEBEIVMHSFNCGEFFYCNSDTLPNST 296  
 QY 343 TNNNTGSSNTEGNTITLPCRIKQIINMMOEVGKAMVAPPIRGQIRCSSNTITGLLTRDG 402  
 DB 297 TNNNTGSSNTEGNTITLPCRIKQIINMMOEVGKAMVAPPIRGQIRCSSNTITGLLTRDG 356  
 QY 403 GGIN-ENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 462  
 DB 357 GGIN-ENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 416  
 QY 463 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQRMQLTVMGIGKOLA 522  
 DB 417 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQRMQLTVMGIGKOLA 476  
 QY 523 QARVAVERYLGDQOOLGIGWCSGKLICTTAVPNNASWSKSLDRINNNMTMEMEREID 582  
 DB 477 QARVAVERYLGDQOOLGIGWCSGKLICTTAVPNNASWSKSLDRINNNMTMEMEREID 536  
 QY 583 NYTSEIYTLIEESONQOEKNEOELLELDKVASLMMWFDITKWLWY 625  
 DB 537 NYTSEIYTLIEESONQOEKNEOELLELDKVASLMMWFDITKWLWY 579

RESULT 9  
 US-10-026-741-103  
 ; Sequence 103, Application US/10026741  
 ; Publication No. US20030049604A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHARNEAU, PIERRE  
 ; APPLICANT: CLAVEL, FRANCOISE  
 ; APPLICANT: BORMAN, ANDREW  
 ; APPLICANT: OULLENT, CAROLINE  
 ; APPLICANT: GUETARD, DENISE  
 ; APPLICANT: MONTAGNER, LUC  
 ; APPLICANT: COHEN, JACQUES  
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR SUBTYPE) ANTIGENS  
 ; NUMBER OF SEQUENCES: 103  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.  
 ; STREET: 1300 I Street, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; APPLICATION NUMBER: US/10/026,741  
 ; FILING DATE: 27-Dec-2001

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/817,441  
 FILING DATE: 31-AUG-1998  
 APPLICATION NUMBER: PCT/FR 95/01391  
 FILING DATE: 20-OCT-1995  
 APPLICATION NUMBER: FR 9412554  
 FILING DATE: 20-OCT-1994  
 APPLICATION NUMBER: FR 9502526  
 FILING DATE: 03-MAR-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meyers, Kenneth J.  
 REGISTRATION NUMBER: 25,146  
 REFERENCE/DOCKET NUMBER: 03260, 6005-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4400  
 INFORMATION FOR SEQ ID NO: 103:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 861 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
 US-10-026-741-103

Query Match 86.6%; Score 2901.5; DB 15; Length 861;  
 Best Local Similarity 84.3%; Pred. No. 3.9e-268; Indels 31; Gaps 7;  
 Matches 552; Conservative 31; Mismatches 41;

2 EKLWVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 61  
 32 EKLWVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 91  
 62 HFNWKNWVQWQEDDIISLMDQSLKPCVKLTPLCVTLNCKDVAATNTDSEG- 115  
 92 NFNWKNWVQWQEDDIISLMDQSLKPCVKLTPLCVTLNCKDVAATNTDSEG 151  
 116 -TMRGEIKNGSPNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 174  
 152 MMEGELKNGSPNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 211  
 175 KISEPPIPIHACAPGAFALCKNDKTPNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLA 234  
 212 KVSPEPIPIHACAPGAFALCKNDKTPNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLA 271  
 235 EEEVVISRDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275  
 272 EEEVVISRDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 331  
 276 RQACNINIRAKMNDTLKQIVIKLREOF-ENKTIIVFNHSSGDPPEIVMHSFNGGGEFFYCN 334  
 332 RQACNINIRAKMNDTLKQIVIKLREOF-ENKTIIVFNHSSGDPPEIVMHSFNGGGEFFYCN 391  
 335 STQFNSTWNTSTEGSNTTETTLPCRIKQIIMMOQEVKAMVAPPIRQIRCS 390  
 392 STQFNSTWNTSTEGSNTTETTLPCRIKQIIMMOQEVKAMVAPPIRQIRCS 451  
 391 SNITGLLTTRDGINENGTIFRPGGDMRDWMSSELYKYKVKIEPLGVAFTKAKRKYV 450  
 452 SNITGLLTTRDGINENGTIFRPGGDMRDWMSSELYKYKVKIEPLGVAFTKAKRKYV 511  
 451 QREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAOQHL 510  
 512 QREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAOQHL 571  
 511 QLTWVGICQLQARVLAVERYLGDQOLGIGWGSGLICTTAVPMNASNSKSLDRIMNM 570  
 572 QLTWVGICQLQARVLAVERYLGDQOLGIGWGSGLICTTAVPMNASNSKSLDRIMNM 631  
 571 TWMEEREIDNTSYTYLIEESONQOEKNEOELLELDKMASLWMPNTITWLMY 625  
 632 TWMEEREIDNTSYTYLIEESONQOEKNEOELLELDKMASLWMPNTITWLMY 686

RESULT 10  
 US-09-476-242-1  
 Sequence 1, Application US/09476242  
 Patent No. US20020146683A1  
 GENERAL INFORMATION:  
 APPLICANT: BARNETT, Susan  
 APPLICANT: HARTOG, Karin  
 APPLICANT: MARTIN, Eric  
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
 FILE REFERENCE: 1605.002  
 CURRENT APPLICATION NUMBER: US/09/476,242  
 CURRENT FILING DATE: 1999-12-30  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 1  
 LENGTH: 856  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus  
 US-09-476-242-1

Query Match 85.9%; Score 2878; DB 10; Length 856;  
 Best Local Similarity 83.7%; Pred. No. 6.8e-266; Indels 26; Gaps 5;  
 Matches 544; Conservative 35; Mismatches 45;

2 EKLWVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 61  
 32 EKLWVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 91  
 62 HFNWKNWVQWQEDDIISLMDQSLKPCVKLTPLCVTLNCKDVAATNTDSEG- 115  
 92 NFNWKNWVQWQEDDIISLMDQSLKPCVKLTPLCVTLNCKDVAATNTDSEG 151  
 120 GEINCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 179  
 152 GEINCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 211  
 180 PIPHYCAPGAFALCKNDKTPNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLAEEV 239  
 212 PIPHYCAPGAFALCKNDKTPNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLAEEV 271  
 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 280  
 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 331  
 281 NISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGDPPEIVMHSFNGGGEFFYCN 339  
 332 NISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGDPPEIVMHSFNGGGEFFYCN 391  
 340 NSTWNN--NTEGSNTEG-NTITLPCRIKQIIMMOQEVKAMVAPPIRQIRCS 395  
 392 NSTWNN--NTEGSNTEG-NTITLPCRIKQIIMMOQEVKAMVAPPIRQIRCS 451  
 396 LLLTRDGINENGTIFRPGGDMRDWMSSELYKYKVKIEPLGVAFTKAKRKYV 455  
 452 LLLTRDGINENGTIFRPGGDMRDWMSSELYKYKVKIEPLGVAFTKAKRKYV 511  
 456 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAOQHL 515  
 512 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAOQHL 571  
 516 GIKQOLQARVLAVERYLGDQOLGIGWGSGLICTTAVPMNASNSKSLDRIMNM 575  
 572 GIKQOLQARVLAVERYLGDQOLGIGWGSGLICTTAVPMNASNSKSLDRIMNM 631  
 576 EREIDNTSYTYLIEESONQOEKNEOELLELDKMASLWMPNTITWLMY 625  
 632 EREIDNTSYTYLIEESONQOEKNEOELLELDKMASLWMPNTITWLMY 681

RESULT 11  
 US-10-196-515-11

```

; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celja C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

```

```

Query Match      85.6%; Score 2867; DB 15; Length 856;
Best Local Similarity 83.5%; Pred. No. 7,7e-265;
Matches 543; Conservative 35; Mismatches 46; Indels 26; Gaps 5;

QY 2 EKLWTVVYGVVWVWEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENYTE 61
DB 32 EKLWTVVYGVVWVWEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENYTE 91
QY 62 HNNMKNMNVEMOEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 92 NFDNMKNMNVEMOEDIIISLDQSLKPCVKLTPLCVSLCTDCLKNDTNTSSSGRMINEK 151
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVXVINDNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEVOKEYALFYKLDIIPIDNPTSYLLISCDTSVITQACPKISFE 211
QY 160 PIPHYCAPAGFAIILKNDKTFNGKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 239
DB 212 PIPHYCAPAGFAIILKNDKTFNGKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 271
QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNN-----GAGDIRQAHG 280
DB 272 IRSVFTDNAKTIIVQLNTSVINCTRPNNTRKRIIRIHRGGRAFVTVGKIGNRQAHG 331
QY 281 NISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSSGDDPEIVHSHFNCGEFFYCNSSTOLF 339
DB 332 NISRAKMDTLKQIVIKLREOFENKTIIFKQSSGDDPEIVHSHFNCGEFFYCNSSTOLF 391
QY 340 NSTWNN--NTWGSNNTEG-NTITLPCRKIQIIMMOEVGKAMYAPRIGQIRCSSNTG 395
DB 392 NSTWNN--NTWGSNNTEG-NTITLPCRKIQIIMMOEVGKAMYAPRIGQIRCSSNTG 451
QY 396 LLTRDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAPTKARRVQREKRAV 455
DB 452 LLTRDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAPTKARRVQREKRAV 511
QY 456 AVGIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTVW 515
DB 512 AVGIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTVW 571
QY 516 GIKQIARVLAVERYLGDQOLIGWCSGKLICTTAVPWNASWNSKSLDRIMNNTWME 575
DB 572 GIKQIARVLAVERYLGDQOLIGWCSGKLICTTAVPWNASWNSKSLDRIMNNTWME 631
QY 576 EREDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNFDITTKMLY 625
DB 632 EREDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNFDITTKMLY 681

```

```

RESULT 12
US-10-196-515-3
; Sequence 3, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celja C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

```

```

Query Match      84.2%; Score 2822.5; DB 15; Length 726;
Best Local Similarity 82.4%; Pred. No. 1.1e-260;
Matches 533; Conservative 38; Mismatches 51; Indels 25; Gaps 5;

QY 2 EKLWTVVYGVVWVWEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENYTE 61
DB 32 EKLWTVVYGVVWVWEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENYTE 91
QY 62 HNNMKNMNVEMOEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 92 NFDNMKNMNVEMOEDIIISLDQSLKPCVKLTPLCVSLCTDCLKNDTNTSSSGRMINEK 151
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVXVINDNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEVOKEYALFYKLDIIPIDNPTSYLLISCDTSVITQACPKISFE 211
QY 160 PIPHYCAPAGFAIILKNDKTFNGKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 239
DB 212 PIPHYCAPAGFAIILKNDKTFNGKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 271
QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNN-----GAGDIRQAHG 280
DB 272 IRSVFTDNAKTIIVQLNTSVINCTRPNNTRKRIIRIHRGGRAFVTVGKIGNRQAHG 331
QY 281 NISRAKMDTLKQIVIKLREOF-ENKTIIVFNHSSGDDPEIVHSHFNCGEFFYCNSSTOLF 339
DB 332 NISRAKMDTLKQIVIKLREOFENKTIIFKQSSGDDPEIVHSHFNCGEFFYCNSSTOLF 391
QY 340 NSTWNN--NTWGSNNTEG-NTITLPCRKIQIIMMOEVGKAMYAPRIGQIRCSSNTG 398
DB 392 NSTW--NTWGSNNTEG-NTITLPCRKIQIIMMOEVGKAMYAPRIGQIRCSSNTG 449
QY 396 LLTRDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAPTKARRVQREKRAV 458
DB 452 LLTRDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAPTKARRVQREKRAV 509
QY 456 AVGIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTVW 518
DB 512 AVGIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLRAIEAOQRLQLTVW 579
QY 516 GIKQIARVLAVERYLGDQOLIGWCSGKLICTTAVPWNASWNSKSLDRIMNNTWME 578
DB 572 GIKQIARVLAVERYLGDQOLIGWCSGKLICTTAVPWNASWNSKSLDRIMNNTWME 629
QY 576 EREDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNFDITTKMLY 625
DB 632 EREDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNFDITTKMLY 676

```

RESULT 13  
US-10-196-515-12  
Sequence 12, Application US/10196515  
Publication No. US20030091594A1  
GENERAL INFORMATION:  
APPLICANT: LABRANCHE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
TITLE OF INVENTION: THERAPEUTICS  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/10/196,515  
CURRENT FILING DATE: 2002-07-16  
PRIOR APPLICATION NUMBER: US/09/337,387  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 12  
LENGTH: 759  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-12

Query Match 84.0%; Score 2815.5; DB 15; Length 759;  
Best Local Similarity 82.4%; Pred. No. 5.4e-260;  
Matches 533; Conservative 38; Mismatches 51; Indels 25; Gaps 5;

QY 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTBEH 61  
DB 32 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTBEH 91  
QY 62 HFNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 119  
DB 92 HFNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 151  
QY 120 GEINCKSFNTTTSIRDEYQKEVALFYKLDVYXIDNNNTSYLISCDTSVITQACPISFE 179  
DB 152 GEINCKSFNTTTSIRDEYQKEVALFYKLDVYXIDNNNTSYLISCDTSVITQACPISFE 211  
QY 180 PIPHYCAPAGFALIKNDKTENGKXPKKNVSTXOCTHGIRPVVSTOLLNGSLAEEBV 239  
DB 212 PIPHYCAPAGFALIKNDKTENGKXPKKNVSTXOCTHGIRPVVSTOLLNGSLAEEBV 271  
QY 240 IRSNFTNNAKTIIIVQLKESVEINCTRENN-----GAGDIRQAH 280  
DB 272 IRSNFTNNAKTIIIVQLKESVEINCTRENN-----GAGDIRQAH 331  
QY 281 NISAKKNDITKQIVIKREOF-ENKTVFNHSSGGDEIYMHSGGGEFFYCNSTOLF 339  
DB 332 NISAKKNDITKQIVIKREOF-ENKTVFNHSSGGDEIYMHSGGGEFFYCNSTOLF 391  
QY 340 NSTNNNTGSGNTEG--NTITLPCRKQIINMOWEKGAMVAPPIRGQIRCSSNITGLL 398  
DB 392 NST--STKSGNNTGSGNTEG--NTITLPCRKQIINMOWEKGAMVAPPIRGQIRCSSNITGLL 449  
QY 399 TRDGINENGTEIFRPGGDMRDNMRSSELYKXKVVKIEPLGVAFTKARRVVOKEKAVG 458  
DB 450 TRDGINENGTEIFRPGGDMRDNMRSSELYKXKVVKIEPLGVAFTKARRVVOKEKAVG 509  
QY 459 IGAVFLGLGAAGSTMGAASMTLTVQARLLISGIVQOQNNLLRAIEAOQRLQLTWGIX 518  
DB 510 IGAVFLGLGAAGSTMGAASMTLTVQARLLISGIVQOQNNLLRAIEAOQRLQLTWGIX 569  
QY 519 QLOARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSKSLDRIINNMWMEWRE 578  
DB 570 QLOARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSKSLDRIINNMWMEWRE 629  
QY 579 IDNTSEIYTLIESONQOEKNEBELLELDKMASIMNFDITKMLWY 625

DB 630 INNTSLIHSLEBSQWQEMNEBELLELDKMASIMNFDITKMLWY 676

RESULT 14  
US-09-938-406-1  
Sequence 1, Application US/09938406  
Patent No. US20020155120A1  
GENERAL INFORMATION:  
APPLICANT: Vancolt, George  
APPLICANT: Vancolt, Thomas  
APPLICANT: Birt, Deborah  
TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR  
TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY  
FILE REFERENCE: 40646-20002.10  
CURRENT APPLICATION NUMBER: US/09/938,406  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 09/214,701  
PRIOR FILING DATE: 1999-09-30  
PRIOR APPLICATION NUMBER: PCT/US 97/12253  
PRIOR FILING DATE: 1997-07-10  
PRIOR APPLICATION NUMBER: US 60/021,687  
PRIOR FILING DATE: 1996-07-10  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 868  
TYPE: PRT  
ORGANISM: Virus HIV-1  
US-09-938-406-1

Query Match 83.4%; Score 2793.5; DB 10; Length 868;  
Best Local Similarity 80.1%; Pred. No. 8.4e-258;  
Matches 528; Conservative 36; Mismatches 58; Indels 37; Gaps 5;

QY 4 LMTVTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTBEH 63  
DB 35 LMTVTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTBEH 94  
QY 64 HFNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 117  
DB 95 HFNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 154  
QY 118 GEINCKSFNTTTSIRDEYQKEVALFYKLDVYXIDNNNTSYLISCDTSVITQ 171  
DB 155 GEINCKSFNTTTSIRDEYQKEVALFYKLDVYXIDNNNTSYLISCDTSVITQ 214  
QY 172 ACPRISFPIPIHYCAPAGFALIKNDKTENGKXPKKNVSTXOCTHGIRPVVSTOLLNG 231  
DB 215 ACPRISFPIPIHYCAPAGFALIKNDKTENGKXPKKNVSTXOCTHGIRPVVSTOLLNG 274  
QY 232 SLAEEVVISDFTNNAKTIIIVQLKESVEINCTRENN-----AG 273  
DB 275 SLAEEVVISDFTNNAKTIIIVQLKESVEINCTRENN-----AG 334  
QY 274 DIRQAHNISRANKNDITKQIVIKREOF-ENKTVFNHSSGGDEIYMHSGGGEFFYCNSTOLF 333  
DB 335 DIRQAHNISRANKNDITKQIVIKREOF-ENKTVFNHSSGGDEIYMHSGGGEFFYCNSTOLF 394  
QY 334 NSTQLFNSTNNNTGSGNTEG--NTITLPCRKQIINMOWEKGAMVAPPIRGQIRCSSNITGLL 387  
DB 395 NSTQLFNSTNNNTGSGNTEG--NTITLPCRKQIINMOWEKGAMVAPPIRGQIRCSSNITGLL 449  
QY 388 RCSNNTIGLLTRDGINENGTEIFRPGGDMRDNMRSSELYKXKVVKIEPLGVAFTKARR 447  
DB 455 RCSNNTIGLLTRDGINENGTEIFRPGGDMRDNMRSSELYKXKVVKIEPLGVAFTKARR 514  
QY 448 RVVQREKRAVG-IGAVFLGLGAAGSTMGAASMTLTVQARLLISGIVQOQNNLLRAIEAO 506  
DB 515 RVVQREKRAVG-IGAVFLGLGAAGSTMGAASMTLTVQARLLISGIVQOQNNLLRAIEAO 574  
QY 507 QRLQLTWGIXIQLQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSKSLDRI 566

Db 575 QHLLQUTWGCIGLOARILAVERYLKDQQLGFWGSGKLICTTAVPWNASNSKTLDOI 634  
 QY 567 WNNMTWMEWERIDNTSEIYTLIEESONQOEKNEQELLELDKMSLWMPDITKWLWY 625  
 Db 635 WNNMTWMEWRIDNTSEIYTLIEESONQOEKNEQELLELDKMSLWMPDITKWLWY 633

Db 573 NYTNLIYTLIEESONQOEKNEQELLELDKMSLWMPDISKWLWY 617  
 Search completed: December 12, 2003, 12:45:24  
 Job time : 27.0558 secs

# RESULT 15

US-09-891-609-4  
 ; Sequence 4, Application US/09891609  
 ; Patent No. US20020127238A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stematacos, Leonidia  
 ; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
 ; FILE REFERENCE: 2570-1-002N  
 ; CURRENT APPLICATION NUMBER: US/09/891,609  
 ; PRIOR FILING DATE: 2001-06-26  
 ; PRIOR APPLICATION NUMBER: 60/214,608  
 ; PRIOR FILING DATE: 2000-06-27  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 619  
 ; TYPE: PRT  
 ; ORGANISM: Human immunodeficiency virus type 1  
 US-09-891-609-4

Query Match 82.3%; Score 2758; DB 10; Length 619;  
 Best Local Similarity 82.0%; Pred. No. 1.3e-254;  
 Matches 529; Conservative 24; Mismatches 42; Indels 50; Gaps 7;

QY 1 VEKLTWTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60  
 Db 3 VEKLTWTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 62  
 QY 61 EHFNMKNMVEQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDV.NAINTNDSGTMER 119  
 Db 63 ENFMNMKNMVEQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDV.NAINTNDSGTMER 122  
 QY 120 GEIKNCSFNTTISINDEVOKEALPYKLDVVAIDNNNTSYRLISCDTSVITQACPKISFE 179  
 Db 123 GEIKNCSFNVGAG-----KLNCTSVITQACPKVSFE 155  
 QY 180 PPIHYCAPAGFAILKNDKTFNGKXPCQNVSTXOCTHGIRPVSTOLLNGSLAEEVY 239  
 Db 156 PPIHYCAPAGFAILKNDKTFNGSGPCTNVSTVOCTHGIRPVSTOLLNGSLAEEVY 215  
 QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHGN 281  
 Db 216 IRSENPFDNAKTIIVQLKESVEINCTRPNNNTRKSTITIGPRAFYATGDIIGDIRQAHGN 275  
 QY 282 ISRAKNDTLKOIVIKLREQFNKTIIVFNHSSGGDEIYMSFNGCGEFFYCNSTOLFNS 341  
 Db 276 ISGEKKNNTLKOIVIKLREQFNKTIIVFNHSSGGDEIYMSFNGCGEFFYCNSTOLFNS 335  
 QY 342 TWNNNTGSGNTEGNTITLPCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLLTRD 401  
 Db 336 TW-NTTIGPNTNG-TITLPCRIKQIINNMOEYKAMVAPPIRGQIRCSNITGLLTRD 393  
 QY 402 GGIN-ENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRRVVOREKRAVIG 460  
 Db 394 GGEKISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRRVVOREKRAVIG 453  
 QY 461 AVFLGFLGAAGSTGAASMTLTVQARLLISGIVQOONNLLRAIEAQRMLOLTVWGIK-L 520  
 Db 454 AMFLGFLGAAGSTGAASMTLTVQARLLISGIVQOONNLLRAIEAQRMLOLTVWGIK-L 512  
 QY 521 QARVLAVERYLGDQQLGIMGSGKLICTTAVPWNASNSKSLDRINNMTWMEWERID 580  
 Db 513 QARVLAVERYLKDQQLGIMGSGKLICTTAVPWNASNSKSLDRINNMTWMEWERID 572  
 QY 581 NYTSEIYTLIEESONQOEKNEQELLELDKMSLWMPDITKWLWY 625  
 Db 573 NYTSEIYTLIEESONQOEKNEQELLELDKMSLWMPDITKWLWY 617

**THIS PAGE BLANK (US:TD)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:25:38 / Search time 10.49 Seconds

(Without alignments)  
2801,881 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351

Sequence: 1 VEKLMVTYGVGVWKEATT.....ELDKWASIMNEDITKMLMY 625

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

1 number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3008.5	89.8	848	1	ENV_HV1JR
2	2928	87.4	847	1	ENV_HV1SI
3	2903.5	86.6	856	1	ENV_HV1SC
4	2901.5	86.6	861	1	ENV_HV1BR
5	2887	86.2	856	1	ENV_HV1BI
6	2884	86.1	867	1	ENV_HV1J3
7	2883	86.0	856	1	ENV_HV1PV
8	2882	86.0	843	1	ENV_HV1Y2
9	2878	85.9	852	1	ENV_HV1BN
10	2878	85.9	856	1	ENV_HV1H2
11	2871	85.7	856	1	ENV_HV1H3
12	2870	85.6	856	1	ENV_HV1MN
13	2864	85.5	856	1	ENV_HV1LW
14	2861.5	85.4	851	1	ENV_HV1B8
15	2861.5	85.4	856	1	ENV_HV1W3
16	2845.5	84.9	852	1	ENV_HV1A3
17	2836	84.6	855	1	ENV_HV1A2
18	2830	84.5	853	1	ENV_HV1MP
19	2818	84.1	847	1	ENV_HV1W2
20	2817	84.1	855	1	ENV_HV1YR
21	2810	83.9	855	1	ENV_HV1OY
22	2793.5	83.4	868	1	ENV_HV1C4
23	2790.5	83.3	861	1	ENV_HV1XB
24	2645	78.9	855	1	ENV_HV1Z6
25	2641	78.8	863	1	ENV_HV1Z8
26	2636	78.7	853	1	ENV_HV1Z2
27	2578.5	76.9	846	1	ENV_HV1MA
28	2572.5	76.8	859	1	ENV_HV1ND
29	2562	75.5	853	1	ENV_HV1ZL
30	2517.5	75.1	856	1	ENV_HV1ZL
31	2180	65.1	854	1	ENV_SIVC2
32	1631.5	48.7	421	1	ENV_HV1N5
33	1542	46.0	460	1	ENV_HV1Z3

34	1199.5	35.8	854	1	ENV_SIVAI	Q02837 simian immu
35	1195.5	35.7	851	1	ENV_HV2G1	P18040 human immu
36	1194	35.6	859	1	ENV_HV2ST	P20872 human immu
37	1176	35.1	712	1	ENV_HV2S2	P32536 human immu
38	1173.5	35.0	768	1	ENV_SIVAI	P27757 simian immu
39	1165	34.8	877	1	ENV_SIVAG	P27977 simian immu
40	1164	34.7	858	1	ENV_HV2RO	P04577 human immu
41	1158	34.6	865	1	ENV_SIVAT	P05805 simian immu
42	1148.5	34.3	859	1	ENV_HV2CA	P24105 human immu
43	1143	34.1	856	1	ENV_HV2NZ	P05883 human immu
44	1136	33.9	882	1	ENV_SIVM1	P05885 simian immu
45	1135.5	33.9	851	1	ENV_HV2D1	P17755 human immu

## ALIGNMENTS

RESULT 1	ENV_HV1JR	STANDARD;	PRT;	848 AA.
AC	P20871;			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV			
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).			
OC	Viruses; Retroviridae; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11688;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Koyanagi S., Chen I.S.Y.;			
RL	Submitted (DEC-1988) to the HIV data bank.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
DR	EMBL: M38429; AAB03749.1; -			
DR	PDB: 1CE4; 1G-MAR-99.			
DR	HIV; M38429; ENV5URCSF.			
DR	InterPro; IPR000328; Env_GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;			
KW	3D-structure.			
FT	SIGNAL	1	32	
FT	CHAIN	33	503	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	848	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73	BY SIMILARITY.
FT	DISULFID	118	203	BY SIMILARITY.
FT	DISULFID	125	194	BY SIMILARITY.
FT	DISULFID	130	154	BY SIMILARITY.
FT	DISULFID	216	245	BY SIMILARITY.
FT	DISULFID	226	237	BY SIMILARITY.
FT	DISULFID	294	328	BY SIMILARITY.
FT	DISULFID	374	437	BY SIMILARITY.
FT	DISULFID	381	410	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	137	133	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	153	157	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	157	157	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	228	228	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .)

```

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 848 AA; 96475 MW; 2076751227EC3F3 CRC64;

Query Match 89.8%; Score 3008.5; DB 1; Length 848;
Best Local Similarity 88.2%; Pred. No. 3.2e-232;
Matches 569; Conservative 22; Mismatches 33; Indels 21; Gaps 4;

QY 1 VEKLVTVYVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVDPNFOEVLENTV 60
DB 30 VEKLVTVYVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVDPNFOEVLENTV 89
QY 61 EHFMMKNNWVQOEDTISLMDOSLKPCKVLTPLCVTLNCKDVNATNTNDSCTMERG 120
DB 90 EDFMMKNNWVQOEDVIMLMDOSLKPCKVLTPLCVTLNCKDVNATNTNDSCTMERG 149
QY 121 EIKKCSFNTISIDENQKEXALPKYLDVXIDN-NNTSYLLISGDSVITQACPKISFE 179
DB 150 EIKKCSFNTISIDENQKEXALPKYLDVXIDN-NNTSYLLISGDSVITQACPKISFE 209
QY 180 PIPHYCAPAFALIKNDKTFNGKPKCKNVTXOCTGIRPVYSTOLLNGSLAESENV 239
DB 210 PIPHYCAPAFALIKNDKTFNGKPKCKNVTXOCTGIRPVYSTOLLNGSLAESENV 269
QY 240 IRSNFTNNAKTITIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281
DB 270 IRSNFTNNAKTITIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 329
QY 282 ISRAKNDTLKOIYIKLEOFENKTIVPNHSSGGPEIYVMSFNGGFEFFCNSTOLFNS 341
DB 330 ISRAKNDTLKOIYIKLEOFENKTIVPNHSSGGPEIYVMSFNGGFEFFCNSTOLFNS 389
QY 342 TWNNNTGSGNNTGEG-TTLPCRIKQIINMNOEVKAYAPPIRQOICSSNITGLLTLR 400
DB 390 TW-NDTESSGTEGNDITILPCRIKQIINMNOEVKAYAPPIRQOICSSNITGLLTLR 448
QY 401 DGGINENGEIIFRPGGDMRDNMWSELKYKVVXIEPLGVAPTKCKRRVORERKAVGIG 460
DB 449 DGGKNESEIEIFRPGGDMRDNMWSELKYKVVXIEPLGVAPTKCKRRVORERKAVGIG 508
QY 461 AVPLGFGAAGSTGASMTLTVQARLLSGIVQOONNLLAIREQOQMLDLTWGIGIOL 520
DB 509 ALPLGFGAAGSTGASMTLTVQARLLSGIVQOONNLLAIREQOQMLDLTWGIGIOL 568
QY 521 QARVLAVERVYLDQOOLGIMGCSGKLICTAVPNMASNSKSLDIRINNMNTMWEVEREID 580
DB 569 QARVLAVERVYLDQOOLGIMGCSGKLICTAVPNMASNSKSLDIRINNMNTMWEVEREID 628
QY 581 NYTSEIYTLIEESQNOQKNEOELLELDKMASLWMPDITKWLWY 625
DB 629 NYTSEIYTLIEESQNOQKNEOELLELDKMASLWMPDITKWLWY 673

```

```

AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirinae.
NCBI_TaxID=11691;
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quirroz M., Tung J.W., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism: cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M65024; AAA5072.1; -.
CC PDB: 1OBE; 15-MAY-97.
CC DR HIV; M38428; ENV$SF162.
CC DR InterPro; IPR000328; Env_GP41.
CC DR InterPro; IPR000777; GP120.
CC DR Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; polyprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 29
CC CHAIN 30 502
CC FT DIULFID 503 847
CC FT DIULFID 53 73
CC FT DIULFID 118 203
CC FT DIULFID 125 194
CC FT DIULFID 130 155
CC FT DIULFID 216 245
CC FT DIULFID 226 237
CC FT DIULFID 294 328
CC FT DIULFID 374 435
CC FT DIULFID 381 408
CC FT CARBOHYD 87
CC FT CARBOHYD 135 135
CC FT CARBOHYD 154 154
CC FT CARBOHYD 186 186
CC FT CARBOHYD 195 195
CC FT CARBOHYD 232 232
CC FT CARBOHYD 239 239
CC FT CARBOHYD 260 260
CC FT CARBOHYD 274 274
CC FT CARBOHYD 293 293
CC FT CARBOHYD 299 299
CC FT CARBOHYD 329 329
CC FT CARBOHYD 336 336
CC FT CARBOHYD 352 352
CC FT CARBOHYD 382 382
CC FT CARBOHYD 388 388
CC FT CARBOHYD 392 392
CC FT CARBOHYD 398 398
CC FT CARBOHYD 401 401
CC FT CARBOHYD 438 438
CC FT CARBOHYD 454 454
CC FT CARBOHYD 602 602
CC FT CARBOHYD 607 607
CC FT CARBOHYD 616 616
CC FT CARBOHYD 628 628
CC SEQUENCE 847 AA; 96135 MW; 0A901317DF7F2AB CRC64;

```



Query Match Similarity 87.4% Score 2928; DB 1; Length 847;  
Best Local Similarity 86.0%; Pred. No. 8,5e-226;  
Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5

QY 1 VEKLVATVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60  
DB 30 VEKLVATVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89  
QY 61 EHFMMKKMMWQEQMEDIIISLDQSLKPCVKLTPLCTVLNCKDV-NAYTTNDSBGTMR 119  
DB 90 ENFMNMKNMWEQMEHDIISLDQSLKPCVKLTPLCTVLNCKNLKNAATKSSNMWEMDR 149  
QY 120 GEIKNCSFNITTSIDVEQKVALPYKLDVXYIDNNNTSYRLISCDTSVITQACPSISE 179  
DB 150 GEIKNCSKRVTTSSINKKQKEXALPYKLDVPEIDNNNTSYKLINCTSVITQACPSVSE 209  
QY 180 PIPHYCAPAGFAILKCNCKTGNKCKPCKNVSTXOCTHGIRPVSTOLLNSLAEVAV 239  
DB 210 PIPHYCAPAGFAILKCNCKKNGSGPCNTVSTVQCTHGRVVSQTOLLNSLAEVAV 269  
QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHCN 281  
DB 270 IRSENFNTNNAKTIIVQLKESVEINCTRPNNNRKSIITIGPGRAFYATGDIIDIRQAHCN 329  
QY 282 ISRAKMNLTQKQIVKLEQFENKTIIVFHSSGSGPEIYMSFNGCGEFCYNSTOLFMS 341  
DB 330 ISGEKMNLTQKQIVKLDQFENKTIIVFQSSGSGPEIYMSFNGCGEFCYNSTOLFMS 389  
QY 342 TWNNNTSGSNNTGNTITLPCRKQIIMNMQEVGRAMYPPIRGQIRCSNITGLLTND 401  
DB 390 TW-NNTIGPNNNG--TITLPCRKQIINRMQEVGRAMYPPIRGQIRCSNITGLLTND 447  
QY 402 GGIN-ENGTEIFRPQCGDMRDNRMSELYKYKVKIEPLGVAPTCKKRRVYQREKRAVGIG 460  
DB 448 GGEKISNTEIFRPQCGDMRDNRMSELYKYKVKIEPLGVAPTCKKRRVYQREKRAVTIG 507  
QY 461 AVPLGFLGAGSTWGAASWTLTVQARLLLSGIVQOQNNLTBAIIEAQRMLOLTWNGIKOL 520  
DB 508 AMPLGFLGAGSTWGAASWTLTVQARLLLSGIVQOQNNLTBAIIEAQRMLOLTWNGIKOL 567  
QY 521 QARVLAVERYLGDQQLIGIWGSGKLICTTAVPMNASWNSKSLDIRIMNMWMEWERID 580  
DB 568 QARVLAVERYLKDQQLIGIWGSGKLICTTAVPMNASWNSKSLDQIMNMWMEWERID 627  
QY 581 NYTSEIYTLIEBSNQOEKNEBELLEDKMASLWPFDTTKWLMY 625  
DB 628 NYTMIYTLIEBSNQOEKNEBELLEDKMASLWPFDTTKWLMY 672

RESULT 3  
ENV\_HVASC STANDARD; PRT; 856 AA.  
AC P05878;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-JUL-1993 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120), Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirinae.  
OX NCBI\_TaxID=11702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88219542; PubMed=3369091;  
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti B., Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;  
RT "Envelope sequences of two new United States HIV-1 isolates.",  
RL Virology 164:531-536(1988).  
CC -I- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN  
1984 IN SOUTHERN CALIFORNIA.

[illegible]

```

Db 91 NFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCVTLNCLNLTNDSTNATNTSSNRG 150
Qy 116 TMRGEIKNCSFNTTTSIRDEVOKEVALFKYLDVYXIDNNNTSRLISCSOTSVTQACPK 175
Db 151 KMEGEMTNCSPNTTTSIRSVQKEVALFKYLDVYDID--NTSTYLLNCLNNTSVTQACPK 208
Qy 176 ISEPEPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTXQCTGIRPVYSTOLLNGLSLAE 235
Db 209 VSEPEPIHYHCA-RWFAIILNCKKFGNGTGPCTNVSTVQCTHGIRPVYSTOLLNGLSLAE 267
Qy 236 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNGA-----GDIRQ 277
Db 268 EEEVIRSENFDTNNAKTIIVOLKESVEINCTRPNNNTTSIHIGPRAFYATGDIIGDIRQ 327
Qy 278 AHNCISAKNNKDTLKOVIKLRBOEFKNTVFNHSSGGDEPIVHNSNCGSEFFYCNSTQ 337
Db 328 AHNCISAKNNKNTLKOVIKLRDQFENKTIIFNNSGGDEPIVHNSNCGSEFFYCNSTQ 387
Qy 338 LFNSTNNNTGNSNTGNTGN--TTLPCRKQIINNMOEGKAMVAPRIGQIRCSNITGL 396
Db 388 LFSSTW-NGTBSNNTGNDITLPCRKKEIINNMOEGKAMVAPRIGQIRCSNITGL 446
Qy 397 LITRDGGINENG-----TEIFRGGGDMRDNRSELYKXVVKIEPLGVAFTKARRVQ 451
Db 447 LITRDGNSKXGSKNENTEIFRPGGDMRDNRSELYKXVVKIEPLGVAFTKARRVQ 506
Qy 452 REKRVAG-IGAVPLGFGAAGSTMGASMTLTVOARLLSGIYOOONNLRAIQAQRM 510
Db 507 REKRAVGTIGAMFGFGAAGSTMGASMTLTVOARLLSGIYOOONNLRAIQAQRM 566
Qy 511 QLTVMGIKQIOLARVLAVERVYLGDOQLLGWCSGKLICTTAVPNNASMSNKSIDRINNM 570
Db 567 QLTVMGIKQIOLARVLAVERVYLRDQQLLGWCSGKLICTTAVPNTSMSNKSIDKINGM 626
Qy 571 TMMEMEREIDNYTSEIITLIESONQOKNEQELLEDKASLNNWEDITKMLY 625
Db 627 TMMEMEREIDNYTSLIYTLIESONQOKNEQELLEDKASLNNWEDITKMLY 681

RESULT 4
ENV_HV1BR STANDARD; PRT; 861 AA.
ID ENV_HV1BR STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=8509333; PubMed=2981635;
RA Main-Hobson S., Sonigo P., Danos O., Cole S., Allion M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
-----
DR EMBL: K02013; AAB59751.1; -
DR EMBL: A04321; CA00352.1; -
DR PIR: A03975; VCLJLV.
DR PDB: 1ERF; 20-FEB-02.
DR HIV: K02013; ENVSBRU.
DR InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR PIR: A03975; GP41; 1.
DR AIDS: Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516
FT CHAIN 517 861
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 453 453
FT CARBOHYD 468 468
FT CARBOHYD 616 616
FT CARBOHYD 621 621
FT CARBOHYD 630 630
FT CARBOHYD 642 642
FT CARBOHYD 679 679
FT CARBOHYD 755 755
FT CARBOHYD 821 821
SQ SEQUENCE 861 AA; 97487 MW; 04DE2BAD4EAFD63A CRC64;

Query Match 86.6%; Score 2901.5; DB 1; Length 861;
Best Local Similarity 84.3%; Pred. No. 1.1e-223;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

Qy 2 EKLMTVTVYGVVPWKATTTLPFCASDAKAYDTEVHNMVATHACVPTDPNPOEVLEVTE 61
Db 32 EKLMTVTVYGVVPWKATTTLPFCASDAKAYDTEVHNMVATHACVPTDPNPOEVLEVTE 91

Qy 62 HFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCVTLNCLNLTNDSTNATNTSSNRG 115
Db 92 NFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCVTLNCLNLTNDSTNATNTSSNRG 151

Qy 116 TMRGEIKNCSFNTTTSIRDEVOKEVALFKYLDVYXIDNNNTSRLISCSOTSVTQACPK 174
Db 152 KMEGEMTNCSPNTTTSIRSVQKEVALFKYLDVYDID--NTSTYLLNCLNNTSVTQACPK 211

Qy 176 ISEPEPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTXQCTGIRPVYSTOLLNGLSLAE 234
Db 212 KMEGEMTNCSPNTTTSIRSVQKEVALFKYLDVYDID--NTSTYLLNCLNNTSVTQACPK 271

Qy 236 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNGA-----GDIRQ 275
Db 268 EEEVIRSENFDTNNAKTIIVOLKESVEINCTRPNNNTTSIHIGPRAFYATGDIIGDIRQ 327

```

```

Db 272 EEEVIRSANFTDNKTTIIVOLNQSVENCTRPNNTRKSIIRIORGPGRAFTIGKIGNM 331
Qy 276 ROAHGNISRAKNADTLKQIVILAREOF-ENKTIIVFNHSGSDPEIYHMSFNGCGEFPYCN 334
Db 332 ROAHGNISRAKNADTLKQIVILAREOF-ENKTIIVFNHSGSDPEIYHMSFNGCGEFPYCN 391
Qy 335 STOLFNSWTWN--NTEGSNNTG-NTITLPCRIOIIMMOEYKAMYPPIRQOIRCS 390
Db 392 STOLFNSWTWNSTGSGNTEGSDTITLPCRIOIIMMOEYKAMYPPIRQOIRCS 451
Qy 391 SNITGILLTRDGGIENGTEIFRPGGDMRWSELYKYKYKIEPLGVAATCKRRV 450
Db 452 SNITGILLTRDGGIENGTEIFRPGGDMRWSELYKYKYKIEPLGVAATCKRRV 511
Qy 451 OREKRAVGIGAVFLGFLGAGSTMGASWTLTVQARLLISGIVQOONNLRAIEAQOHL 510
Db 512 OREKRAVGIGAVFLGFLGAGSTMGASWTLTVQARLLISGIVQOONNLRAIEAQOHL 571
Qy 571 QLTWGIKOLQARVLAVERYLGDQQLGIMGCSGKLICTTAVPMNASMSNKSIDRIWMNM 570
Db 572 QLTWGIKOLQARVLAVERYLGDQQLGIMGCSGKLICTTAVPMNASMSNKSIDRIWMNM 631
Qy 571 TMMEBERIDNTSITITLIESONQOEKNEBELLDKMAWLPDITTKLWY 625
Db 632 TMMEBERIDNTSITITLIESONQOEKNEBELLDKMAWLPDITTKLWY 686

```

## RESULT 5

```

ENV_HV1B1
ID ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;

```

```

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
CX NCBI_TaxID=11678;
RN [1]
RP MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Habeline W., Patarca R., Liyak K.J., Starcich B.R.,
  Joseph S.F., Doran B.R., Rafalski J.A., Whitehorn B.A.,
  Baumeister K., Ivanoff L., Peteway S.R., Pearson M.L.,
  Lautenberger J.A., Papas T.S., Grayed J., Chang N.T., Gallo R.C.,
  Wong-Staal F.;
  "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
  Nature 313:277-284 (1985).

```

```

RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
  Gregory T.J.;
  "Assignment of intrachain disulfide bonds and characterization of
  potential glycosylation sites of the type 1 recombinant human
  immunodeficiency virus envelope glycoprotein (gp120) expressed in
  Chinese hamster ovary cells.";
  J. Biol. Chem. 265:10373-10382 (1990).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC EMBL: M15654; AAA44205.1;
DR PIR: A03973; VCLRH3.
DR HIV: M15654; ENV5BH102.
DR InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
  Signal.

```

```

FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 286 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 625 625
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

```

```

Query Match 86.2%; Score 2887; DB 1; Length 856;
Beet Local Similarity 84.0%; Pred. No. 1.6e-222;
Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

```

```

Qy 2 EKLWTVVYGVVWMEATTTTLCASDAKYTEVAVNVAHACVPTDPNPOEVLENTTE 61
Db 32 EKLWTVVYGVVWMEATTTTLCASDAKYTEVAVNVAHACVPTDPNPOEVLENTTE 91
Qy 62 HFNWKNMNVQEDDIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
Db 92 HFNWKNMNVQEDDIISLDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 151
Qy 120 GEIKKCSNITTSIRDBQKRYALFYKLDVXKIDNNNTSYLLISGDTSVITQACPKISFE 179
Db 152 GEIKKCSNITTSIRDBQKRYALFYKLDVXKIDNNNTSYLLISGDTSVITQACPKISFE 211
Qy 180 PIPHYCAPAGFALIKCNDKTFNGCKPKKNVSTXQCTGIRPVVSTOLLNGLSLAESEV 239
Db 212 PIPHYCAPAGFALIKCNDKTFNGCKPKKNVSTXQCTGIRPVVSTOLLNGLSLAESEV 271
Qy 240 IRSDFNNAKTIIVOLKESVEINCTRPNN--GAGDIRQAH 280
Db 272 IRSANFTDNKTTIIVOLNQSVENCTRPNNTRKSIIRIORGPGRAFTIGKIGNRQAH 331

```

QY 281 NISRAKNDTLKOIVIKLREOF-ENKTIIVFNHSSGDPPEIYVMSHSGGFEFFYCNSTOLF 339  
 DB 332 NISRAKNNLTKQDLSKREDFGNKNTIIFKSSGSGDEPIYTHSFGNGGEFFYCNSTOLF 391  
 QY 340 NSTWNN--NTEGSNTEG-NTITLPCRIOIIMMOEYGVKAWAPPRIQIRCSSNITG 395  
 DB 392 NSTWNNSTWSTKSGNTEGSDITLPCRIOIIMMOEYGVKAWAPPRIQIRCSSNITG 451  
 QY 396 LLTRDGIINENGTEIFRPGGDMRDMWRSLEYKYVKIIEPLGVAATKCRVRVOREKR 455  
 DB 452 LLTRDGSNNSEIFRPGGDMRDMWRSLEYKYVKIIEPLGVAATKCRVRVOREKR 511  
 QY 456 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQRLQTLTW 515  
 DB 512 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAQRLQTLTW 571  
 QY 516 GIKQARVLAVERYLDDQQLIGWCSGKLICTTAVPWNASWNSKSLDRIMNMTWMEW 575  
 DB 572 GIKQARVLAVERYLDDQQLIGWCSGKLICTTAVPWNASWNSKSLDRIMNMTWMEW 631  
 QY 576 EREIDNTSEIYTLIESONOEKNEOBLFELDKWASLWMPDITKMLY 625  
 DB 632 EREIDNTSEIYTLIESONOEKNEOBLFELDKWASLWMPDITKMLY 681  
 RESULT 6  
 ENV\_HV1J3 STANDARD; PRT; 867 AA.  
 ID ENV\_HV1J3  
 AC P12489;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).  
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
 NC NCBI\_Taxid=11694;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9352108; PubMed=2669897;  
 RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;  
 RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria.";  
 RT AIDS Res. Hum. Retroviruses 5:411-419(1989).  
 RL AIDS Res. Hum. Retroviruses 5:411-419(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M21138; AAB03526.1; -  
 DR HIV; M21138; ENV:JH3.  
 DR InterPro: IPR000328; Env GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KM AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
 KW signal.  
 FT SIGNAL. 1 30  
 FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 217 BY SIMILARITY.  
 FT DISULFID 125 208 BY SIMILARITY.  
 FT DISULFID 130 160 BY SIMILARITY.  
 FT DISULFID 230 259 BY SIMILARITY.  
 FT DISULFID 240 251 BY SIMILARITY.  
 FT DISULFID 308 342 BY SIMILARITY.

FT DISULFID 388 457 BY SIMILARITY.  
 FT DISULFID 395 430 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 867 AA; 98399 MW; 5F23101468B86860 CRC64;  
 Query Match 86.1%; Score 2884; DB 1; Length 867;  
 Best Local Similarity 83.6%; Pred. No. 2,9e-222;  
 Matches 554; Conservativity 25; Mismatches 44; Indels 40; Gaps 6;  
 QY 2 EKLMTVYGVVPWKEATTLTFCASDAKAYDTEVHANVATHACVPDPNPOEVLEVTE 61  
 DB 31 EQLMTVYGVVPWKEATTLTFCASDAKAYDTEVHANVATHACVPDPNPOEVLEVTE 90  
 QY 62 HFNWKKNNWEOQOEDIIISLMDQSLKPCVLTPLCVTLNCKD-----VNATNTDSEG 115  
 DB 91 KFNWKKNNWEOQOEDIIISLMDQSLKPCVLTPLCVTLNCKD-----VNATNTDSEG 150  
 QY 116 TMRGEIKNSFNITTSIRDEVOKEVALFYKLDVYXIDN-----NNTSRLISCDT 166  
 DB 151 KMEGEMKNOSFNITTSIRKVKKEHALFYKHVDPINNSTKONIKNDSTRRLISCDT 210  
 QY 167 SVTQACPKISPEPIPIHYCAPGAFALKGNKDTFNGKXPKNAVSTQCTHGRPVVSTQ 226  
 DB 211 SVTQACPKISPEPIPIHYCAPGAFALKGNKDTFNGKXPKNAVSTQCTHGRPVVSTQ 270  
 QY 227 LLNGLSLAESEVYIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG----- 271  
 DB 271 LLNGLSLAESEVYIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG----- 330  
 QY 272 ---AGDIRQAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDPPEIYVMSHSGG 328  
 DB 331 KQIAGDIRQAHCNINRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDPPEIYVMSHSGG 390  
 QY 329 EFPYCNSTOLFNSNWTLSNSTWSTKSGNTEGSDITLPCRIOIIMMOEYGVKAWAPP 382  
 DB 391 EFPYCNSTOLFNSNWTLSNSTWSTKSGNTEGSDITLPCRIOIIMMOEYGVKAWAPP 450  
 QY 383 IRGQIRCSSNITGLLTRDGIINENGTEIFRPGGDMRDMWRSLEYKYVKIIEPLGVAP 442  
 DB 451 IREGIRCSSNITGLLTRDGIINENGTEIFRPGGDMRDMWRSLEYKYVKIIEPLGVAP 510  
 QY 443 TKCRVRVOREKRAVGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRA 502  
 DB 511 TKCRVRVOREKRAVGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRA 569

FT	CARBOHYD	230	230	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	234	234	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	241	241	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	262	262	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	276	276	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	289	289	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	295	295	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	301	301	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	332	332	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	339	339	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	356	356	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	386	386	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	392	392	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	397	397	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	406	406	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	448	448	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	463	463	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	611	611	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	616	616	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	625	625	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	637	637	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	674	674	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	750	750	N-LINKED (GLCNAC .)	(POTENTIAL)	
FT	CARBOHYD	816	816	N-LINKED (GLCNAC .)	(POTENTIAL)	
SEQ	SEQUENCE	856 AA;	97339 MW;	5FCD81DC3C1209B3 CRC64;		
Query Match		86.0%;	Score 2883;	DB 1;	Length 856;	
Best local Similarity		83.8%;	Pred. No. 3,46-222;			
Matches	545;	Conservative .33;	Mismatches 46;	Indels 26;	Gaps 5	
QY	2	EKLMTVYVGVVWKEATTTTLFCASDAAYTTEVHNWATHACPTDPNQEVLNVTE	61			
DB	32	EKLMTVYVGVVWKEATTTTLFCASDAAYTTEVHNWATHACPTDPNQEVLNVTE	91			
QY	62	HEMMKNNVNEQMODIISLMDOSLKPVCVKLTPLCTYLKCDVNAITTTDSRG--TMR	119			
DB	92	NFMKMKNDVNEQMHEDIISLMDOSLKPVCVKLTPLCTYLKCDVNAITTTDSRG--TMR	151			
QY	120	GEIKNCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISPE	179			
DB	152	GEIKNCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISPE	211			
QY	180	PIPIHCAAGAILKCKDKTPNGKPKRNVSTYQCTHGIRPVVYSTOLLNGLSABEVV	239			
DB	212	PIPIHCAAGAILKCKDKTPNGKPKRNVSTYQCTHGIRPVVYSTOLLNGLSABEVV	271			
QY	240	INSNDNTNAKTIIVQLKESVEINCTPRNNN-----GAGDIQAH	280			
DB	272	INSANFTDAKTIIVQLKESVEINCTPRNNNTRKSIIRIQGPGRAFTYICKIGNMQAH	311			
QY	281	NISRAKMNLTLOKQIVYLKEQF--ENKTIVPFNHSSGSGPEIVMHSFNGCGEFFYCSNSTOLF	339			
DB	332	NISRAKMNLTLOKQIVYLKEQF--ENKTIVPFNHSSGSGPEIVMHSFNGCGEFFYCSNSTOLF	391			
QY	340	NSTWN--NTEGSNTEG--NTTLPCKIKOIMMOEYGVKAMYPPIRQIRCCSNTG	395			
DB	392	NSTWN--NTEGSNTEG--NTTLPCKIKOIMMOEYGVKAMYPPIRQIRCCSNTG	451			
QY	396	LILLTRDGGINENGTIIFPRGGDMRDNNRSLSLYKYVYKIEPLGVAATKKRVRVOREKX	455			
DB	452	LILLTRDGGINENGTIIFPRGGDMRDNNRSLSLYKYVYKIEPLGVAATKKRVRVOREKX	511			
QY	456	AVGIGAVFLGFLPGAAGSTMGAASTMLTVOALLLSGIVQOONNLLRAIEAQOMLOLTVW	515			
DB	512	AVGIGAVFLGFLPGAAGSTMGAASTMLTVOALLLSGIVQOONNLLRAIEAQOMLOLTVW	571			
QY	516	GIKOLQARLAVERYLADQOLIGMGSGKLICTTAVPMNASMNSKSLDIRINNMNTMMEW	575			
DB	572	GIKOLQARLAVERYLADQOLIGMGSGKLICTTAVPMNASMNSKSLDIRINNMNTMMEW	631			
QY	576	BREINDYNTSEIYTLIEESQONQEKNEDELELDMKASLNNWFDITKMYT	635			
DB	632	BREINDYNTSEIYTLIEESQONQEKNEDELELDMKASLNNWFDITKMYT	691			

FT	CARBOHYD	389	389	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	395	395	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	400	400	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	435	435	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	450	450	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	598	598	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	603	603	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	612	612	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	624	624	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
FT	CARBOHYD	803	803	N-LINKED (GLCNAC . .)	(POTENTIAL) -					
SQ	SEQUENCE	843 AA;	95648 MM;	C69DFD97JC918B71 CRC64;						
 Query Match      86.0%;    Score 2882; DB 1;    Length 843;										
Best Local Similarity    85.1%;    Pred. No. 4e-222;										
Matches    548;    Conservative    28;    Mismatches    42;    Indels    26;    Gaps    7										
Oy	2 EKLWTVVYGVVPMKEATTTLFCASADAKAYDEVHNVNATHACVPTDPNPDEVLNENYTE									
Dd	31 EQLWTVVYGVVPWKAEATTTLFCASADAAYOTEVHNWATHACVPTDPNPQVKLENYTE									
Oy	62 HFNNKNNMVEWOMODIISLMDOSIKPCVKLTPLCTYLNCQDV-NATNTNDSEGMERG									
Dd	91 NFNNKNNMVEWQMEDIISLMDOSIKPCVKLTPLCTVLNCTDLRANATYTSSSWETMEG									
Oy	121 EIKKCSFNTITSIDVOKEYALPYKLDVXYIDNNNSYSRLISCVTSYITQACPISPEP									
Dd	151 EIKKCSFPITTSIDKVOKEYALFNLDLVPID--NASIRLSCTSUYTQAICPVFSFEF									
Oy	181 IPIHYCAPAGFAILKCKNDKTENGKKPKCKNVSXOCTHGIRPVSVSTOLLNGSLAEEVI									
Dd	209 IPIHYCAPAGFAILKCKNDKKENGSGPCINVSIVTOCTHGIRPVSVSTOLLNGSLAEEVI									
Oy	241 RSDNFNNNAKTIIVQLKESVEINCTRPNNNC-----AGDIRAHANI									
Dd	269 RSENFNNNAKTIIVQLNESVVINCTRPNNNRKSINIGGRALYTTGEIIIGDIRAHANI									
Oy	283 SRAKNMDLKIIVLKLRPF--ENKTIIVNHSSGGPEIIVMSHFNGGEPFYCNSQTLEFS									
Dd	329 SKTQMENTLEOIATIKLEQFGNNKTIIFNPSSGGPEIIVTHSPNGGEPFYCNSQTLP--									
Oy	342 TWANNTEGSNTTEGNITFLPCRIOIIMMOEYGKAMAPPIRGQIRCSSNTTGLLTRD									
Dd	387 TW-NDTRKLANIT-GRNTTLPCRIOIIMMOEYGKAMAPPIRGQIRCSSNTTGLLTRD									
Oy	402 GGINENGTEIPRPGGDMDRWRSSELKYKVKIEPLGVAPTCKRRVAVOREKRAVGICA									
Dd	445 GSKDTNGTEIPRPGGDMDRWRSSELKYKVKIEPLGVAAPTAKARRVAREKRAVGIGA									
Oy	462 VFLLGFLGAAGSTMGAASMTLTVQARLLLSGIIVOQNMLLRPAIEAQORMLQLTWMIKIQ									
Dd	505 LFLGFLGAAGSTMGAASITTLTVQARQLLSGIIVOQNMLLRPAIEAOHLLQLTWMIKIQ									
Oy	522 ARVALAVERLYLDQQLLGIGWCGSKLICCTAAPVMAWSNKSIDRIINNNTVMWEEREIDN									
Dd	565 ARVALAVERYLRDQQLLGIGWCGSKLICCTTYPAWTSWNKSLNETIWDMMTMKWEREIDN									
Oy	582 YTSEIYTLIESONQOEKNEQELLELDKVASLMWFDTTKMLMY									
Dd	625 YTHIITYSLIESONQOEKNEQELLALDKVASLMWFDTTKMLMY									
 RESULT 9										
ENV_HV1BN                  STANDARD;                  PRT;        852 AA.										
AC	P12488;									
DT	01-OCT-1989 (Rel. 12, Created)									
DT	01-OCT-1989 (Rel. 12, Last sequence update)									
DT	15-SEP-2003 (Rel. 42, Last annotation update)									
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].									
DE	ENV.									
NN										
OC	Human immunodeficiency virus type 1 (IBR isolate) (HIV-1). Viruses; Retroviral viruses; Retroviridae; Lentivirus.									





SQ	SEQUENCE	8656 AA; 97212 MW; 6FBA16AF65107FE0 C6C64;
	Query Match	85.9%; Score 2878; DB 1; Length 856;
	Best Local Similarity	83.7%; Pred. No. 8.5e-22;
	Matches 544; Conservative 35; Mismatches 45; Indels 26; Gaps 5	
OY	EKLWVTVYGVGPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDENPOEVLNTE	61
Db	32 EKLWVTVYGVGPWKAEATTLTFCASDAKAYDTEVHNWATHACVPTDENPOEVLNTE	91
OY	62 HFNKKNNMVEQOMEDIISLWDQSILKPCVKLTPLCVTLNCKDVANTNTNSEG--TWER	119
Db	92 NFNWMKMDMVEQOMEDIISLWDQSILKPCVKLTPLCVSLKCTDLKNDITNTNSSGHIMBK	151
OY	120 GEIKNCSPNTTSIRDEVOKEALFLKYLDVXIXINNNTSYLLISCDTSVITQACKISFE	179
Db	152 GEIKNCSPNTSIRGRVQKEAFAFFKYLDIPIIDDTITSYLKLSNTSVITQACRVSFE	211
OY	180 PIPHYCAPAGFALKCNKDFTENGXPCKANSTYQCTHGIRPVSTOLLNGSLAESEV	239
Db	212 PIPHYCAPAGFALKCNKNKTFGNGPCTNVSTVOCTHGINPVSTOLLNGSLAESEV	271
OY	240 IRSDNFTNAKTIIVOLKESVEINCTREPENN-----GAGDIRQAHC	280
Db	272 IRSNFTDNAKTIIVQLNTSVEINCTRNNTTRKRIRIQRGRAFAVTIGKIGNRQAHC	331
OY	281 NISPAKNMTLKQIVIKIREGF-EKTIIVPHSSGGDEIVMHSFNGCGEPFYCNSTOLF	339
Db	332 NISPAKNMTLKQIASKIREQFGNNKTIIFQSSGSGDEIVTHSFNCGEPFYCNSTOLF	391
OY	340 NSTANN---NEGSSNNEG-NTITLPGCIKOIIMMOVGKAMYPPIRGQIRGSNTTG	395
Db	392 NSTPNSTWSSTEGSNNEGSDTITLPCKIKOIIIMMQKVGAAMPPIISGLRCSSNTTG	451
OY	396 LLLTRDDGINSNGTEIFRPGCGDMRDNRSELKYKKVXIEPLGYAPTKCRKRVVQREK	455
Db	452 LLLTRDGNSNSNEBEIFRPGGDWRDRNRSSELYKKVXIETLGVAPEPKARKRVVQREK	511
OY	456 AVGIGAVFLGPLGAAGSTMGAASMTLVQAELLISGIYQQONNLRAIEAQRMILTVM	515
Db	512 AVGIGAVFLGPLGAAGSTMGAASMTLVQAQLLSGTIYQQONNLRAIEAOHLILTVM	571
OY	516 GIKLOARVLAVERVYGQQLLGIMGSCGKIICCAVPWMAWSWSKSIDRIINNMNTWME	575
Db	572 GIKLOAAIILAVERLKQQLLGIMGSCGKICTTAVPWMAWSWSKSLEQIINNMNTWME	631
OY	576 EREDINYSEIYTLIESONOENOEELBELDKVASLMMNFDTIKMLMY	625
Db	632 DREINNYTSLHSILIESONOENOEELBELDKVASLMMNFNTIMLMY	681
RESULT 11		
ENV_HV1H3		
ID_ENV HV1H3	STANDARD; PRt; 856 AA.	
AC	P04624;	
DC	13-AUG-1987 (Rel. 05, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].	
GN	ENV.	
OS	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).	
OC	Virusess; Retroid viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11707;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=8528248; PubMed=2988795;	
RA	Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R., Shaw G.M., Wong-Staal F., Reddy E.P.;	
RT	"HIV-III env gene products synthesized in E. coli are recognized by antibodies present in the sera of AIDS patients.";	
RL	Cell 41:979-986(1985).	
CC	-----	





DR PDB; INJO; 25-FEB-03.  
 DR HIV; M17449; ENVSMN.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR007777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;  
 3D-structure.  
 KM SIGNAL 1 29  
 FT CHAIN 30 513 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 514 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 210 BY SIMILARITY.  
 FT DISULFID 125 201 BY SIMILARITY.  
 FT DISULFID 130 162 BY SIMILARITY.  
 FT DISULFID 223 252 BY SIMILARITY.  
 FT DISULFID 233 244 BY SIMILARITY.  
 FT DISULFID 301 335 BY SIMILARITY.  
 FT DISULFID 381 445 BY SIMILARITY.  
 FT DISULFID 388 418 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 626 626 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97140 MW; D197B08940BE732 CRC64;  
 Query Match 85.6%; Score 2870; DB 1; Length 856;  
 Best Local Similarity 83.4%; Pred. No. 3.7e-22;  
 Matches 548; Conservative 29; Mismatches 42; Indels 38; Gaps 7;

QY 234 AAEVVRISDNFTNNAKTIIVOLKESVEINCRPN-----NNGADI 275  
 DB 271 AAEVVRISDNFTNNAKTIIVOLKESVEINCRPNVKRRIHIGPARAYTTNIGTI 330  
 QY 276 ROAHCHISRAKMDTLKQIYIKLREOPENKTIYPNHSGGDPPIVMSFNCGEFFCNS 335  
 DB 331 ROAHCHISRAKMDTLKQIYIKLREOPENKTIYPNHSGGDPPIVMSFNCGEFFCNS 390  
 QY 336 TOLFNSITWN-----NNEGSSNTEGNTITLPCRIKOIINMOEGKMYAPPIRGORCS 390  
 DB 391 SPLFNSITWNNGNTNNNTGSN---NNITLQCKIKIINMOEGKMYAPPIRGORCS 446  
 QY 391 SNITGLLTREDCG--INENGTEIFRPGGDMRDNRSELYKYVVKIIEPGLVATPKCR 448  
 DB 447 SNITGLLTREDCG--INENGTEIFRPGGDMRDNRSELYKYVVKIIEPGLVATPKCR 506  
 QY 449 VVOREKRAVIGAVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEQOR 508  
 DB 507 VVOREKRA-AIGALFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEQOR 565  
 QY 509 MLOLTWGIKOLQARVLAVERVIGDQOLGIMGSGKLICTAVPNVNASMNSKSLDIRWN 568  
 DB 566 MLOLTWGIKOLQARVLAVERVIGDQOLGIMGSGKLICTAVPNVNASMNSKSLDIRWN 625  
 QY 569 NMTWMEEREIDNVTSEIYTLIEBSQNOERKEBELLELDKMASLWNPDIITKLMY 625  
 DB 626 NMTWMEEREIDNVTSEIYTLIEBSQNOERKEBELLELDKMASLWNPDIITKLMY 682

RESULT 13  
 ENV\_HVILW STANDARD; PRT; 856 AA.  
 AC 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane  
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (TM12.3 isolate) (HIV-1).  
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
 CC NCBI\_TaxID=82834;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95127297; PubMed=7826699;  
 RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
 Shaw G.M., Kong L.T., Weiss S.H., Waters D., Gallo R.C., Blattner W.,  
 "Viral variability and serum antibody response in a laboratory worker  
 infected with HIV type 1 (HIV type 11B).";  
 RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; J12053; AA76690.1; -  
 DR PDB; 1IF3; 02-MAY-01.  
 DR GLYCOSULEDB; Q70626; -  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR007777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;  
 3D-structure.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 54 74 BY SIMILARITY.  
 FT DISULFID 119 205 BY SIMILARITY.

```

FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 96938 MW; 0C241332CF7E6687 CRC64;

```

```

Query Match 85.5%; Score 2864; DB 1; Length 856;
Best Local Similarity 83.5%; Pred. No. 1,le-20;
Matches 543; Conservative 36; Mismatches 45; Indels 26; Gaps 6;

Oy 2 EKLAWTVYVGVPMVEATTTTLCASDAKAYDREVNWATHACVTPDDNPOCVLENTE 61
Db 32 EKLAWTVYVGVPMVEATTTTLCASDAKAYDREVNWATHACVTPDDNPOCVLENTE 91
62 HFNMMKNNMVEQMOEDISLMDQSLKPCVKLTPLCVTLNCKOV-NATNTTNDSEG-TMER 119
Db 92 NFNMMKNNMVEQMOEDISLMDQSLKPCVKLTPLCVTLNCKOV-NATNTTNDSEG-TMER 151
Oy 120 GEIKKCSNITTSIRDEVQKEAYALFYKLDVXIDNNNTSYRLISCDSVTITQACPKISFE 179
Db 152 GEIKKCSNITTSIRDEVQKEAYALFYKLDVXIDNNNTSYRLISCDSVTITQACPKISFE 211
Oy 180 PIPHYCAPAGPAILKCKDKTFNGKXPKCNVSTACTGIRPVSTOLLNGSLAEEVV 239
Db 212 PIPHYCAPAGPAILKCKDKTFNGKXPKCNVSTACTGIRPVSTOLLNGSLAEEVV 271
Oy 240 IRSNDFTNNAKTIIVQLKESVEINCTRPNN-----GAGDIRQAHG 280
Db 272 IRSNDFTNNAKTIIVQLKESVEINCTRPNN-----GAGDIRQAHG 331
Oy 281 NISPAKNDTLKQIVIKLRQEP-ENKTVFNHSSGADPEIVMHSFNGCEFFYCNSTOLF 339
Db 332 NISPAKNDTLKQIVIKLRQEP-ENKTVFNHSSGADPEIVMHSFNGCEFFYCNSTOLF 391
Oy 340 NSTNNN--NTEGSNTEG-NITTLPCRIKQIINNMVGVKAMVAPRIKQIRCSNITG 395
Db 392 NSTNNN--NTEGSNTEG-NITTLPCRIKQIINNMVGVKAMVAPRIKQIRCSNITG 451
Oy 396 LLLTRDGGINRNGTEIFRRPGGDMRDNMRSELKYKVKVIEPLGVAFTPKCRKRVVQREKR 455

```

```

Db 452 LLLTRDGGINRNGTEIFRRPGGDMRDNMRSELKYKVKVIEPLGVAFTPKCRKRVVQREKR 511
Oy 456 AVIGAVELGLGAAGSMGAASMTLTQVQALLSGIVQOQNNILRAEAOQRMQLTVW 515
Db 512 AVIGAVELGLGAAGSMGAASMTLTQVQALLSGIVQOQNNILRAEAOQRMQLTVW 571
Oy 516 GIKOLQARVLAVERLDGQQLIGWCSGSKLICCTAVPWNASWNSKSIDRTMNNNTMMEW 575
Db 572 GIKOLQARVLAVERLDGQQLIGWCSGSKLICCTAVPWNASWNSKSIDRTMNNNTMMEW 631
Oy 576 EREIDNTSYETTLIESONQOEKNEQELBELDKRASLMMNPDTKWLWY 625
Db 632 DREINNTYSLHSLIESONQOEKNEQELBELDKRASLMMNPDTKWLWY 681

RESULT 14
ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
ID 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein Gp160 precursor [Contains: Exterior membrane
GN glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
OS Human immunodeficiency virus type 1 (B8 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
NC NCBT_TaxID=11684;
RX MEDLINE=8511123; PubMed=2578615;
RA Joseph S.F., Dotan E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Pelletier S.R., Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk.)
CC -----
CC EMBL: K02011; AAA4461.1; -.
CC PDB: 1A1K; 16-JUN-97.
CC PDB: 1DDH; 13-JAN-99.
CC PDB: 1HHG; 31-OCT-93.
CC PDB: 1QO3; 02-JAN-00.
CC PDB: 1S2T; 24-DEC-97.
CC HIV: K02011; ENVSBH8.
CC Glycosylated: P04582; -.
CC InterPro: IPR000328; Env_Gp41.
CC InterPro: IPR000777; Gp120.
CC Pfam: PF00516; Gp120; 1.
CC AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 30
CC CHAIN 31 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
CC CHAIN 507 851 TRANSMEMBRANE GLYCOPROTEIN.
CC DISULFID 54 74 BY SIMILARITY.
CC DISULFID 119 205 BY SIMILARITY.
CC DISULFID 126 196 BY SIMILARITY.
CC DISULFID 131 157 BY SIMILARITY.
CC DISULFID 218 247 BY SIMILARITY.
CC DISULFID 228 239 BY SIMILARITY.
CC DISULFID 296 331 BY SIMILARITY.
CC DISULFID 378 445 BY SIMILARITY.

```

	FT	DISULFID	385	413	BY SIMILARITY.
	FT	CARBOHYD	88	88	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	136	136	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	141	141	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	156	156	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	160	160	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	186	186	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	197	197	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	230	230	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	234	234	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	241	241	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	262	262	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	276	276	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	295	295	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	301	301	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	332	332	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	339	339	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	356	356	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	386	386	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	392	392	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	401	401	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	443	443	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	458	458	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	606	606	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	611	611	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	620	620	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	632	632	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	669	669	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	745	745	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	FT	CARBOHYD	811	811	N-LINKED (GLCNAC . . .) (POTENTIAL) .
	SEQ	SEQUENCE	851 AA;	96644 MW;	D16A3C9085785F1 CRC64;
	Query Match		85.4%;	Score 2861.5;	DB 1; Length 851;
	Best Local Similarity		83.8%;	Pred. No. 1.8e-220;	
	Matches 542;	Conservative	33;	Mismatches 47;	Indels 25; Gaps 5;
Qy	2	EKLMTVTVYGVYVMEATTTTLEFCASDARAVIDEVHNVNATHACVPTDPPQOEVLLENTE	61		
Db	32	EKLMTVTVYGVYVMEATTTTLEFCASDARAVIDEVHNVNATHACVPTDPPQOEVLLENTE	91		
Qy	62	HNFMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDQVNAINTNTNSEG--TWER	119		
Db	92	NFMKNNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDQVNAINTNTNSEG--TWER	151		
Qy	120	GEIKKCSFNITTSIDDEVQKEYALVYKLDVXKIDNNNTSYRLISCDTSTYTTQACPKISFE	179		
Db	152	GEIKKCSFNITTSIDDEVQKEYALVYKLDVXKIDNNNTSYRLISCDTSTYTTQACPKISFE	211		
Qy	180	PIPIHYCAPAGAILLKNDKTFNGKXCKPCNVSTQCTGIRPVNSTOILLNSLAEFEV	239		
Db	212	PIPIHYCAPAGAILLKNDKTFNGKXCKPCNVSTQCTGIRPVNSTOILLNSLAEFEV	271		
Qy	240	IRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIROAH	280		
Db	272	IRSVNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIROAH	331		
Qy	281	NISRAKMDTLKQIVIKLREOF-EKTTIVFNHSGSDPEIVMHSFNCGGEFFYCNSTOLF	339		
Db	332	NISRAKMDTLKQIVIKLREOF-EKTTIVFNHSGSDPEIVMHSFNCGGEFFYCNSTOLF	391		
Qy	340	NSTNNNNEGSNNTEG-NITTLPCGIIKOIINNMQVGVKMAPPIRGQIRGSNTTGLL	398		
Db	392	NSTNNNNEGSNNTEG-NITTLPCGIIKOIINNMQVGVKMAPPIRGQIRGSNTTGLL	449		
Qy	399	TRDGSINENGTETFRPGGDMRDNMRSELVKKVVKIEBLGAPFKCKARVVQREKRAVG	458		
Db	450	TRDGSINENGTETFRPGGDMRDNMRSELVKKVVKIEBLGAPFKCKARVVQREKRAVG	509		
Qy	459	IGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOQNNLLRAIQAQRMQLQTVWGIX	518		
Db	510	IGAVFLGFLGAAGSTWGAASMTLVQARLLLSGIVQOQNNLLRAIQAQRMQLQTVWGIX	569		
Qy	519	QLOARVLAVERLGGQOQLLGTCGCGKGLCTCAVPMNMSKSIDRIINNNTTWMEWE	578		

[illegible]

FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 447 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 SQ SEQUENCE 856 AA; 97526 MW; DB68D1B9C404DE9 CRC64;

Query Match 85.4%; Score 2861.5; DB 1; Length 856;  
 Best Local Similarity 83.6%; Pred. No. 1.8e-220;  
 Matches 547; Conservative 29; Mismatches 47; Indels 31; Gaps 6;

Qy 1 VEKLWTVYVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60  
 Db 30 VEQLVWTVYVYGVVWKEATTTTLFCASDAKAYSTEAKWATHACVPTDNPQEVLENT 89  
 Qy 61 EHFNMKNMNVQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDVANATNTTNDSEGTMERG 120  
 Db 90 EHFNMKNMNVQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDVANATNTTNDSEGTMERG 147  
 Qy 121 EIKNCSFNITTSIRDEQKEVALFKLDVYXIDNNN-----TSYRLISCDTSVITQACP 175  
 Db 148 EYKNCSEFNITTSIRKRVKEXVALFKLDVYVPIKSNDSSTYRRLHCHNTSVITQACSK 207  
 Qy 176 ISFEPPIHYCAPAGFALIKCNDKTENGCKPCKNVSTXOCTHGIRPVVSTQLLNGSLAE 235  
 Db 208 VSFEPPIHYCAPAGFALIKCNDKFFNGTGPCNTVSTVQCTHGIRPVVSTQLLNGSLAE 267  
 Qy 236 EEVTVRSDFNTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQ 277  
 Db 268 EEVTVRSDFNTNNAKTIIVQLKESVEINCTRPNNNVRRRHHIGPRAFYTGIRGINIRO 327  
 Qy 278 AHCNISRAKMDTLKOIVIKLAEOPENKTIVFNHSSGGDEIWMHSFNGGSEFFYCNSTQ 337  
 Db 328 AHCNISRAKMDTLKOIVIKLAEOPENKTIVFNHSSGGDEIWMHSFNGGSEFFYCNSTQ 387  
 Qy 338 LFNSTWN--NNTGSSNNT--GNTITLPCRIKOIINMQEYKAMVAPPPIRGQIRCSN 392  
 Db 388 LFNSTWNVTGISTEENNTGENDITLPCRIKOIINMQEYKAMVAPPPIRGQIRCSN 447  
 Qy 393 ITGLLLTRDGINENGTEIFRPGGDMRDNRSELYKYVVKI EPLGVAPTKCKRRVQR 452  
 Db 448 ITGLLLTRDGINSSREIFRPGGDMRDNRSELYKYVVKI EPLGVAPTKCKRRVQR 507  
 Qy 453 EKRANG-IGAVFLGFLGAAGSTMGAASMTLTVQARLLSGIVQOQNNLLRAIEAQQMLQ 511  
 Db 508 EKRANGAIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIQOQNNLLRAIEAQQMLQ 567  
 Qy 512 LTVNGIKQIQAARVLAVERYLADQOLLGIGGSGKLIICCTAVPMNAMSNSKSLDRIMNMT 571  
 Db 568 LTVNGIKQIQAARVLAVERYLADQOLLGIGGSGKLIICCTAVPMNAMSNSKSLDRIMNMT 627  
 Qy 572 WMEMEREIDNYSSEIYTLIEESQNOQEKNOELLEIDKVASLNNWFDITKMLMY 625  
 Db 628 WMEMEREIDNYSSEIYTLIEESQNOQEKNOELLEIDKVASLNNWFDITKMLMY 681

Search completed: December 12, 2003, 12:30:50  
 Job time : 12.49 secs

**THIS PAGE BLANK (08-70)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:27:29 ; Search time 14.5506 Seconds  
(without alignments)  
4130.780 Million cell updates/sec

Title: US-10-032-162-17  
Perfect score: 3351  
Sequence: 1 VEKLMVTYVYGVPMWEKAT.....ELDKWASLMMNFDITKMLMY 625

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	98.3	847	2 T09448	envelope glycoprot
2	3264	97.4	847	2 S13289	env protein - huma
3	2901.5	86.6	861	1 VCLJLV	env polypeptide pr
4	2900.5	86.6	852	1 T12016	envelope glycoprot
5	2887	86.2	856	1 VCLJH3	env polypeptide pr
6	2883	86.0	856	1 VCLJVL	env polypeptide pr
7	2882	86.0	843	1 H44001	env polypeptide pr
8	2881	86.0	854	2 S13288	env protein - huma
9	2875	85.8	852	1 VCLJBR	env polypeptide -
10	2866.5	85.5	851	2 S13985	env polypeptide -
11	2861.5	85.4	851	1 VCLJ3W	env polypeptide pr
12	2859.5	85.3	859	1 VCLJMN	env polypeptide pr
13	2853	85.1	861	1 VCLJSC	env polypeptide pr
14	2836	84.6	855	1 VCLJJC	env polypeptide pr
15	2793.5	83.4	868	1 VCLJH4	env polypeptide -
16	2790.5	83.3	729	1 VCLJXK	env polypeptide pr
17	2790.5	83.3	861	1 VCLJXB	env polypeptide pr
18	2645	78.9	855	1 VCLJZR	env polypeptide pr
19	2636	78.7	853	2 S54384	envelope polype
20	2578.5	76.9	846	1 VCLJND	env polypeptide pr
21	2572.5	76.8	859	2 T01672	envelope polype
22	2517.5	75.1	856	1 A44963	env polypeptide pr
23	2180	65.1	854	1 VCLJSI	env polypeptide pr
24	2058	61.4	506	2 A40218	envelop glycoprote
25	1890.5	56.4	443	2 C41621	env polypeptide P
26	1877.5	56.0	445	2 A41621	env polypeptide M
27	1865.5	55.7	495	2 S31493	env polypeptide -
28	1840	54.9	877	2 S49197	envelope protein P
29	1801	53.7	863	2 A53034	gag polypeptide -

30	1784	53.2	454	2 B41621	env polypeptide D
31	1210.5	36.1	290	2 S25940	env protein - huma
32	1204.5	35.9	297	2 S60538	envelope polype
33	1195.5	35.7	852	1 VCLJG5	env polypeptide pr
34	1195	35.7	859	1 VCLJST	env polypeptide pr
35	1176	35.1	712	1 VCLJH4	env polypeptide pr
36	1165	34.8	877	2 C46356	env polypeptide pr
37	1162	34.7	858	1 VCLJG2	env polypeptide pr
38	1158	34.6	864	1 VCLJG4	env polypeptide pr
39	1148.5	34.3	859	1 VCLJCT	env polypeptide pr
40	1138	34.0	732	2 S46352	env polypeptide -
41	1136	33.9	881	1 VCLJG3	env polypeptide -
42	1135.5	33.9	851	2 S12159	env protein - huma
43	1135.5	33.9	869	2 A47665	env protein gp120 (
44	1134.5	33.9	859	2 S24571	env protein - huma
45	1134	33.8	869	2 S53038	envelope polype

ALIGNMENTS

QY	1	VEKLMVTYVYGVPMWEKATTTLPFCASDAKADTEVHNMATACVPTDPNPOEVLENT	60
DB	30	VEKLMVTYVYGVPMWEKATTTLPFCASDAKADTEVHNMATACVPTDPNPOEVLENT	89
QY	61	EHFMKNNMVEQOEDIIISLMDSLKPCVLTPLCVTLNCKDVNATNTNDSCTMERG	120
DB	90	EHFMKNNMVEQOEDIIISLMDSLKPCVLTPLCVTLNCKDVNATNTNDSCTMERG	149
QY	121	EIKKCSFNTTTSIDDEVQEXALFYKLDVYKIDNNNTSYRLSCTSVITQACPKISFEP	180
DB	150	EIKKCSFNTTTSIDDEVQEXALFYKLDVYKIDNNNTSYRLSCTSVITQACPKISFEP	209
QY	181	IPHYCAPAGFAILKNDKNTNGKPKCNVSTXCTGHRVSVSTOLLNLSLAEVVI	240
DB	210	IPHYCAPAGFAILKNDKNTNGKPKCNVSTXCTGHRVSVSTOLLNLSLAEVVI	269
QY	241	RSDFTNNAKTIIVOLKESVEINCTRPNNNG	282
DB	270	RSDFTNNAKTIIVOLKESVEINCTRPNNNRKSHIGRAFTYTGSHIIDIRAHANI	329
QY	283	SPRAKNDTLKOIVIKLRQFENKTIIVNHSGGDEIYVMSFGCGEFPYCNSTOLFNST	342
DB	310	SPRAKNDTLKOIVIKLRQFENKTIIVNHSGGDEIYVMSFGCGEFPYCNSTOLFNST	389
QY	343	WNNTEGSGNTEGNTITLPCRIKOIINMWSGVKAMVAPPIRGQIRGCSNTTGLLTG	402
DB	390	WNNTEGSGNTEGNTITLPCRIKOIINMWSGVKAMVAPPIRGQIRGCSNTTGLLTG	449
QY	403	GINENGTEIFRPGGDMKDMNRSELYKKYKVKIIEPLGVAPTKCKRRVVRKRAVGIGAV	462

Db 450 GINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGAPTAKARVVOEKAIVGIGAV 509  
 Qy 463 FLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORMLQLTWAGIKOLA 522  
 Db 510 FLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORMLQLTWAGIKOLA 569  
 Qy 523 RVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 582  
 Db 570 RVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 629  
 Qy 583 TSEIYTLIESONOEKNEOELLLEDKWASIMNFDITKMLWY 625  
 Db 630 TSEIYTLIESONOEKNEOELLLEDKWASIMNFDITKMLWY 672

## RESULT 2

env protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
 Accession: S13289  
 R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
 Nature 348, 69-73, 1990  
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
 A:Reference number: S13288; MUID:91043044; PMID:2172833  
 A:Accession: S13289  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-847 <ORF>  
 C:Superfamily: type E retrovirus env polyprotein

Query Match 97.4%; Score 3264; DB 2; Length 847;  
 Best Local Similarity 95.5%; Pred. No. 7,2e-235;  
 Matches 613; Conservative 1; Mismatches 10; Indels 18; Gaps 1;

Qy 2 EKLWVTVYGVVWPKKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 61  
 Db 31 EKLWVTVYGVVWPKKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 90  
 Qy 62 HNNMKNMWEQMEDIIISLDOSLKPCKVLTPLCVTLNCDVNAITVNDSEGTMERGE 121  
 Db 91 HNNMKNMWEQMEDIIISLDOSLKPCKVLTPLCVTLNCDVNAITVNDSEGTMERGE 150  
 Qy 122 IKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSYRLISCDTSVITQACPKISPEPI 181  
 Db 151 IKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSYRLISCDTSVITQACPKISPEPI 210  
 Qy 182 PIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLAEVYIR 241  
 Db 211 PIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLAEVYIR 270  
 Qy 242 SDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGIRQAHCHIS 283  
 Db 271 SDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGBFAVYTGELIGDIRAHCHIS 330  
 Qy 284 RAKNNDTLKOIVILREOFENKTIIVFNHSSGDEPEIWHSPNCGEFPYCNSTOLFNSW 343  
 Db 331 RAKNNDTLKOIVILREOFENKTIIVFNHSSGDEPEIWHSPNCGEFPYCNSTOLFNSW 390  
 Qy 344 NNNTGSSNTEGNTITLPCRIKOIIMNWEVGMAYAPPIRGQIRCSNITGLLTRDG 403  
 Db 391 NNNTGSSNTEGNTITLPCRIKOIIMNWEVGMAYAPPIRGQIRCSNITGLLTRDG 450  
 Qy 404 INENGTETFRPGGDMRDNRSELYKKVYKIEPLGAPTAKARVVOEKAIVGIGAV 463  
 Db 451 INENGTETFRPGGDMRDNRSELYKKVYKIEPLGAPTAKARVVOEKAIVGIGAV 510  
 Qy 464 LGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORMLQLTWAGIKOLA 523  
 Db 511 LGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORMLQLTWAGIKOLA 570  
 Qy 524 VLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 583

Db 571 VLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 630  
 Qy 584 SEIYTLIESONOEKNEOELLLEDKWASIMNFDITKMLWY 625  
 Db 631 SEIYTLIESONOEKNEOELLLEDKWASIMNFDITKMLWY 672

## RESULT 3

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)  
 N:Alternate names: coat polyprotein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
 C:Accession: A03975  
 R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
 Cell 40, 9-17, 1985  
 A:Title: Nucleotide sequence of the AIDS virus, LAV.  
 A:Reference number: A90866; MUID:85093333; PMID:2981635  
 A:Accession: A03975  
 A:Molecule type: DNA  
 A:Residues: 1-861 <NAI>  
 A:Cross-references: GB:K02013; NID:G326417; PIDN:AAB59751.1; PID:G326424  
 C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

Query Match 86.6%; Score 2901.5; DB 1; Length 861;  
 Best Local Similarity 84.3%; Pred. No. 7,2e-208;  
 Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

Qy 2 EKLWVTVYGVVWPKKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 61  
 Db 32 EKLWVTVYGVVWPKKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 91  
 Qy 62 HNNMKNMWEQMEDIIISLDOSLKPCKVLTPLCVTLNCDVNAITVNDSEGTMERGE 115  
 Db 92 HNNMKNMWEQMEDIIISLDOSLKPCKVLTPLCVTLNCDVNAITVNDSEGTMERGE 151  
 Qy 116 -THERGKIKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSYRLISCDTSVITQAC 174  
 Db 152 -THERGKIKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSYRLISCDTSVITQAC 211  
 Qy 175 KISPEPIPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 234  
 Db 212 KISPEPIPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 271  
 Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDI 275  
 Db 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGBFAVYTGELIGDIRAHCHIS 331  
 Qy 276 RQAHCHISRAKNDTLKOIVILREOFENKTIIVFNHSSGDEPEIWHSPNCGEFPYCN 334  
 Db 332 RQAHCHISRAKNDTLKOIVILREOFENKTIIVFNHSSGDEPEIWHSPNCGEFPYCN 391  
 Qy 335 STOLFNSWNTGSSNTEGNTITLPCRIKOIIMNWEVGMAYAPPIRGQIRCSNITGL 390  
 Db 392 STOLFNSWNTGSSNTEGNTITLPCRIKOIIMNWEVGMAYAPPIRGQIRCSNITGL 451  
 Qy 391 SNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGAPTAKARV 450  
 Db 452 SNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGAPTAKARV 511  
 Qy 451 QREKRAVIGAVFLGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORML 510  
 Db 512 QREKRAVIGAVFLGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAEAQORML 571



QY 511 QLTWTCIGTQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNM 570  
 Db 572 QLTWTCIGTQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNM 631  
 QY 571 TMMWERIDNTSYTLIESONQOEKNEBELLEDKMAMNMFDTTKMLY 625  
 Db 632 TMMWERIDNTSYTLIESONQOEKNEBELLEDKMAMNMFDTTKMLY 686

## RESULT 4

T12016

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)

C1Species: human immunodeficiency virus type 1, HIV-1

C1Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C1Accession: T12016

R1McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.

A1Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S

A1Reference number: 217379; PMID:98178716; PMID:9519894

A1Accession: T12016

A1Molecule type: DNA

A1Residues: 1-852 &lt;MCC&gt;

A1Cross-references: EMBL:U90934; NID:G2351783; PIDN:AACS9271.1; PID:G2351784

C1Genetics:

A1Gene: env

C1Superfamily: type E retrovirus env polypeptide

Query Match 86.6%; Score 2900.5; DB 2; Length 852;

Best Local Similarity 85.0%; Pred. No. 8.5e-208;

Matches 550; Conservative 30; Mismatches 44; Indels 23; Gaps 5;

QY 2 EKLWTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61

Db 31 EQLRTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90

QY 62 HNNMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEMR 119

Db 91 NFNMMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEMR 150

QY 120 GEIKKCSFNITTSIRDEYQKAYALFYKLDVYXIDNNNTSYRLISCDTSVITQACKISFE 179

Db 151 GEIKKCSFNITTSIRDEYQKAYALFYKLDVYXIDNNNTSYRLISCDTSVITQACKISFE 210

QY 180 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 239

Db 211 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 270

QY 240 IRSDFNTNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 281

Db 271 IRSDFNTNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 330

QY 282 ISRAKMNNTLKOIYKLBQFENKTIIVFNHSSGDEIVTHSFGCGEFPYCNSTOLF 341

Db 331 ISRAKMNNTLKOIYKLBQFENKTIIVFNHSSGDEIVTHSFGCGEFPYCNSTOLF 390

QY 342 TMM--NNTEGSNTEG--TITLPCRIOKIIINMOEVRGAMVAPPIRGOIRCSNITG 398

Db 391 TMM--NNTEGSNTEG--TITLPCRIOKIIINMOEVRGAMVAPPIRGOIRCSNITG 450

QY 399 TRDGINENGEIIFRPGGDMRDNMRSELKYKVKIIEPLGVAPTKCRKRVVQREKAVG 458

Db 451 TRDGINENGEIIFRPGGDMRDNMRSELKYKVKIIEPLGVAPTKCRKRVVQREKAVG 510

QY 459 IGAFLGLGAGSTMGAAASMTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTWGK 518

Db 511 IGAFLGLGAGSTMGAAASMTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTWGK 570

QY 519 QLOARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTMWERE 578

Db 571 QLOARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTMWERE 630

QY 579 IDNTSYTLIESONQOEKNEBELLEDKMAMNMFDTTKMLY 625

Db 631 IDNTSYTLIESONQOEKNEBELLEDKMAMNMFDTTKMLY 677

## RESULT 5

VCLH3

env polypeptide precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N1Alternate names: coat polypeptide

C1Species: human immunodeficiency virus type 1, HIV-1

A1Note: host Homo sapiens (man)

C1Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C1Accession: A03973

R1Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doran

A1Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A1Reference number: A93353; PMID:85111123; PMID:2578615

A1Accession: A03973

A1Molecule type: DNA

A1Residues: 1-856 &lt;RAT&gt;

A1Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:G326383; PIDN:AAA442

C1Genetics:

A1Gene: env

C1Superfamily: type E retrovirus env polypeptide

C1Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp

F.1-30/Domain: signal sequence #status predicted &lt;SIG&gt;

F.512-856/Product: exterior membrane glycoprotein #status predicted &lt;EXT&gt;

F.188-136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406

F.611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.2%; Score 2887; DB 1; Length 856;

Best Local Similarity 84.0%; Pred. No. 8.6e-207;

Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

QY 2 EKLWTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61

Db 32 EKLWTVYVYGVPMKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91

QY 62 HNNMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEMR 119

Db 92 NFNMMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEMR 151

QY 120 GEIKKCSFNITTSIRDEYQKAYALFYKLDVYXIDNNNTSYRLISCDTSVITQACKISFE 179

Db 152 GEIKKCSFNITTSIRDEYQKAYALFYKLDVYXIDNNNTSYRLISCDTSVITQACKISFE 211

QY 180 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 239

Db 212 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 271

QY 240 IRSDFNTNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 280

Db 272 IRSDFNTNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 331

QY 281 ISRAKMNNTLKOIYKLBQFENKTIIVFNHSSGDEIVTHSFGCGEFPYCNSTOLF 339

Db 332 ISRAKMNNTLKOIYKLBQFENKTIIVFNHSSGDEIVTHSFGCGEFPYCNSTOLF 391

QY 340 NSTWNN--NTEGSNTEG--NTITLPCRIOKIIINMOEVRGAMVAPPIRGOIRCSNITG 395

Db 392 NSTWNN--NTEGSNTEG--NTITLPCRIOKIIINMOEVRGAMVAPPIRGOIRCSNITG 451

QY 396 LLLTRDGINENGEIIFRPGGDMRDNMRSELKYKVKIIEPLGVAPTKCRKRVVQREK 455

Db 452 LLLTRDGINENGEIIFRPGGDMRDNMRSELKYKVKIIEPLGVAPTKCRKRVVQREK 511

QY 456 AVGIGAVFLGLGAGSTMGAAASMTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 515

Db 512 AVGIGAVFLGLGAGSTMGAAASMTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 571

QY 516 GIKOLOARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTMWER 575

Db 572 GIKOLQARILAVERYLKDQQLIGWCSGKLICTTAVPWNASWNSKSLBQIMNMWMEW 631  
 Qy 576 EREIDNTSEIYTLIESQNOEKNEBELLDKMASLWNPDTTKMLY 625  
 Db 632 DREINNTYSLIHSLIESQNOEKNEBELLDKMASLWNPDTTKMLY 681

## RESULT 6

VCL:VTL

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A03974

R:Mesing, M.A.; Smith, D.H.; Cabrerilla, C.D.; Benton, C.V.; Laeky, L.A.; Capon, D.J.

Nature 313, 450-458, 1985

A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi

A:Reference number: A93355; MUID:8511157; PMID:2982104

C:Accession: A03974

Molecule type: DNA

A:Residues: 1-856 &lt;MB&gt;

A:Cross-references: GB:K02083; NID:G555008; PID:AB59873.1; PID:G328559

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

F:1-30/Domain: signal sequence #status predicted &lt;SIG&gt;

F:31-51/Domain: signal sequence #status predicted &lt;SIG&gt;

F:52-856/Product: transmembrane glycoprotein #status predicted &lt;TM&gt;

F:88,136,141,156,160,186,197,230,241,262,276,289,295,301,332,339,356,386,392,397,406

F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.0%; Score 2883; DB 1; Length 856;

Best Local Similarity 83.8%; Pred. No. 1.7e-206;

Matches 545; Conservative 33; Mismatches 46; Indels 26; Gaps 5;

Qy 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 61  
 Db 32 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 91  
 Qy 62 HFNWKNWMEQWEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSEG--TWER 119  
 Db 92 NFNWKNWMEQWEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSEG--TWER 151  
 Qy 120 GEINCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISGDTSVITQACPKISFE 179  
 Db 152 GEINCSFNITTSIRDEVQKEYALFYKLDIIPIDNNTSYRLISGDTSVITQACPKISFE 211  
 Qy 180 PIPHYCAPAGFALIKCDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 239  
 Db 212 PIPHYCAPAGFALIKCDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 271  
 Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAHNC 280  
 Db 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKSIIRQGRGAFVITIGIKIMNRQHC 331  
 Qy 281 NISAKNNDIKQIVIKLEQF-EKTIIVFNHSSGDEPEIVHSHSGGGEFFYCNSDTLF 339  
 Db 332 NISAKNNDIKQIVIKLEQF-EKTIIVFNHSSGDEPEIVHSHSGGGEFFYCNSDTLF 391  
 Qy 340 NSTNNN--NTEGNSNTEG--NTITLPCRIKQIIMMOEYVKAAPPIRGDIRCSSNITG 395  
 Db 392 NSTNNN--NTEGNSNTEG--NTITLPCRIKQIIMMOEYVKAAPPIRGDIRCSSNITG 451  
 Qy 396 LLTRDGGINENGEIIFRPGGDMRDNMRSELKYKVKIIBPLGVAFTKCRVVOREKR 455  
 Db 452 LLTRDGGINENGEIIFRPGGDMRDNMRSELKYKVKIIBPLGVAFTKCRVVOREKR 511  
 Qy 456 AVGIGAVLGLGGAAGSTGAASMTLTVQARLLLSGIVQOONNLIRAEAOQRMQLTIV 515  
 Db 512 AVGIGAVLGLGGAAGSTGAASMTLTVQARLLLSGIVQOONNLIRAEAOQRMQLTIV 571

Qy 516 GIKOLQARILAVERYLKDQQLIGWCSGKLICTTAVPWNASWNSKSLBQIMNMWMEW 575  
 Db 572 GIKOLQARILAVERYLKDQQLIGWCSGKLICTTAVPWNASWNSKSLBQIMNMWMEW 631  
 Qy 576 EREIDNTSEIYTLIESQNOEKNEBELLDKMASLWNPDTTKMLY 625  
 Db 632 DREINNTYSLIHSLIESQNOEKNEBELLDKMASLWNPDTTKMLY 681

## RESULT 7

H44001

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Sep-1994

C:Accession: H44001

R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A:Title: Complete nucleotide sequence, genome organization, and biological properties of

A:Reference number: A44001; MUID:93021387; PMID:1404605

A:Accession: H44001

A:Residues: 1-843 &lt;Li&gt;

A:Cross-references: GB:M93258

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-29/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-35/Region: hydrophobic

F:30-489/Product: coat protein gp120 #status predicted &lt;GP1&gt;

F:490-843/Product: coat protein gp41 #status predicted &lt;GP2&gt;

F:499-515/Region: hydrophobic

F:673-689/Region: hydrophobic

F:738-755/Domain: transmembrane #status predicted &lt;TM&gt;

F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 86.0%; Score 2882; DB 1; Length 843;

Best Local Similarity 85.1%; Pred. No. 2e-206;

Matches 548; Conservative 28; Mismatches 42; Indels 26; Gaps 7;

Qy 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 61  
 Db 31 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENVTE 90  
 Qy 62 HFNWKNWMEQWEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGWERG 120  
 Db 91 NFNWKNWMEQWEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGWERG 150  
 Qy 121 EIKKCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISGDTSVITQACPKISFE 180  
 Db 151 EIKKCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISGDTSVITQACPKISFE 208  
 Qy 181 IPIHYCAPAGFALIKCDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 240  
 Db 209 IPIHYCAPAGFALIKCDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 268  
 Qy 241 RSNFTNNAKTIIVOLKESVEINCTRPNN-----AGDIRQAHNC 282  
 Db 269 RSNFTNNAKTIIVOLKESVEINCTRPNNTRKSIIRQGRGAFVITIGIKIMNRQHC 328  
 Qy 283 SRKANDTLKQIVIKLEQF-EKTIIVFNHSSGDEPEIVHSHSGGGEFFYCNSDTLF 341  
 Db 329 SKTQWETLDEIAIKLEQEGNNKTIIFNPSGGDEPEIVHSHSGGGEFFYCNSDTLF 386  
 Qy 342 TANNNTGSNNTEGNTITLPCRIKQIIMMOEYVKAAPPIRGDIRCSSNITGLTRD 401  
 Db 387 TM-NDTRKLNVT-GRNTTLPCRIKQIIMMOEYVKAAPPIRGDIRCSSNITGLTRD 444  
 Qy 402 GGINENGEIIFRPGGDMRDNMRSELKYKVKIIBPLGVAFTKCRVVOREKRAVIGA 461

Db 445 GGGKDTNGEIEFRPGGDMRDNRSELYKKYKVIIEPLGVAPTKRRVYQREKRAVGLCA 504  
 Qy 462 VFELGFGAAGSTMGASMTLTVOARLLLSGIYQOQNNLRAIEAOQRMQLTWGIGKQ 521  
 Db 505 LFLGLGAGSTMGASMTLTVOARQLSGIYQOQNNLRAIEAOQHLQLTWGIGKQ 564  
 Qy 522 ARVLAVERYLGDQQLLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNNTMMEWEIEDN 581  
 Db 565 ARVLAVERYLRDQQLLGIGWCSGKLICTTAVPMNTSWSNKSINELINDNTMMEWEIEDN 624  
 Qy 582 YTSSEIYTLIEESQNOEKNEQELBLDKASIMNFDITKMLY 625  
 Db 625 YTHIYSLIEESQNOEKNEQELBLDKASIMNFDITKMLY 668

## RESULT 8

env protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 Accession: S13288  
 Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
 A:Reference number: S13288; PMID:91043044; PMID:2172833  
 A:Accession: S13288  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-854 <OBR>  
 C:Superfamily: type E retrovirus env polypeptide

Query Match 86.0%; Score 2881; DB 2; Length 854;  
 Best Local Similarity 84.2%; Pred. No. 2, 4e-206;

Matches 547; Conservative 29; Mismatches 46; Indels 28; Gaps 6;

Qy 2 EKLWTVYVGVVWVKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 61  
 Db 32 EKLWTVYVGVVWVKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 91  
 Qy 62 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCLTLCNKDVNATNTTNSBEG--TNER 119  
 Db 92 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCLTLCNKDVNATNTTNSBEG--TNER 151  
 Qy 120 GEIKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNSYRLISGDSVITQACPKISFE 179  
 Db 152 GEIKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNSYRLISGDSVITQACPKISFE 209  
 Qy 160 PIPHYCAPAGFALIKCNKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNGSLAEEV 239  
 Db 210 PIPHYCAPAGFALIKCNKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNGSLAEEV 269  
 Qy 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 280  
 Db 270 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 329  
 Qy 281 NISRAKNDTLKQIVIKLREOF-ENKTIYFNSSGGDEIVMHSFNGCGEFPYCNSDTOLF 339  
 Db 330 NISRAKNDTLKQIVIKLREOF-ENKTIYFNSSGGDEIVMHSFNGCGEFPYCNSDTOLF 389  
 Qy 340 NSTMNN--NTEGSNNTG--NTITLPCRKQIINMOEYKAMVAPPIRQOICSSNITG 395  
 Db 390 NSTMNN--NTEGSNNTG--NTITLPCRKQIINMOEYKAMVAPPIRQOICSSNITG 449  
 Qy 396 LLTLDGGINENGTIFRRPGGDMRDNRSELYKKYKVIIEPLGVAPTKRRVYQREK 455  
 Db 450 LLTLDGGINENGTIFRRPGGDMRDNRSELYKKYKVIIEPLGVAPTKRRVYQREK 509  
 Qy 456 AVGIGAVFLGFGAAGSTMGASMTLTVOARLLLSGIYQOQNNLRAIEAOQRMQLTW 515  
 Db 510 AVGIGAVFLGFGAAGSTMGASMTLTVOARQLSGIYQOQNNLRAIEAOQHLQLTW 569  
 Qy 516 GIKQLQARVLAVERYLGDQQLLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNNTMME 575

Db 570 GIKQLQARVLAVERYLGDQQLLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNNTMME 629  
 Qy 576 FREINDYTSSEIYTLIEESQNOEKNEQELBLDKASIMNFDITKMLY 625  
 Db 630 FREINDYTSSEIYTLIEESQNOEKNEQELBLDKASIMNFDITKMLY 679

## RESULT 9

env polypeptide - human immunodeficiency virus type 1 (isolate BR)  
 VCLJBR  
 N:Alternate names: coat polypeptide  
 N:Contains: coat protein gp120; coat protein gp41  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: host Homo sapiens (man)  
 C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Feb-1997  
 C:Accession: A31667  
 R:Andand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.  
 Virology 168, 79-89, 1989  
 A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)  
 A:Reference number: A94389; PMID:89085613; PMID:2789516  
 A:Accession: A31667  
 A:Molecule type: DNA  
 A:Residues: 1-852 <ANA>  
 C:Superfamily: type E retrovirus env polypeptide  
 C:Keywords: capsid protein; coat protein; polypeptide; transmembrane protein  
 F:1-516/Product: coat protein gp120 #status predicted <CPI>  
 F:517-852/Product: coat protein gp41 #status predicted <CPI>

Query Match 85.8%; Score 2875; DB 1; Length 852;  
 Best Local Similarity 83.6%; Pred. No. 6, 7e-206;

Matches 541; Conservative 39; Mismatches 43; Indels 24; Gaps 5;

Qy 2 EKLWTVYVGVVWVKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 61  
 Db 32 EKLWTVYVGVVWVKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 91  
 Qy 62 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCLTLCNKDVNATNTTNSBEGMERGE 121  
 Db 92 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCLTLCNKDVNATNTTNSBEGMERGE 151  
 Qy 122 IKNCSEFNITTSIRDEVOKEVALFYKLDVYKIDNN--NTSYRLISGDSVITQACPKISFE 179  
 Db 152 IKNCSEFNITTSIRDEVOKEVALFYKLDVYKIDNN--NTSYRLISGDSVITQACPKISFE 211  
 Qy 160 PIPHYCAPAGFALIKCNKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNGSLAEEV 239  
 Db 212 PIPHYCAPAGFALIKCNKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNGSLAEEV 271  
 Qy 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 281  
 Db 272 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 331  
 Qy 282 ISRAKNDTLKQIVIKLREOF-ENKTIYFNSSGGDEIVMHSFNGCGEFPYCNSDTOLFS 341  
 Db 332 ISRAKNDTLKQIVIKLREOF-ENKTIYFNSSGGDEIVMHSFNGCGEFPYCNSDTOLFS 391  
 Qy 342 TMMNNTBESNTEGNT--ITLPCRKQIINMOEYKAMVAPPIRQOICSSNITLTLTR 400  
 Db 392 TMMNNTBESNTEGNT--ITLPCRKQIINMOEYKAMVAPPIRQOICSSNITLTLTR 450  
 Qy 401 DQGINE--NGTEIFRRPGGDMRDNRSELYKKYKVIIEPLGVAPTKRRVYQREK 458  
 Db 451 DQGINE--NGTEIFRRPGGDMRDNRSELYKKYKVIIEPLGVAPTKRRVYQREK 510  
 Qy 459 IGAVFLGFGAAGSTMGASMTLTVOARLLLSGIYQOQNNLRAIEAOQRMQLTWGIG 518  
 Db 511 IGAVFLGFGAAGSTMGASMTLTVOARLLLSGIYQOQNNLRAIEAOQHLQLTWGIG 570  
 Qy 519 QLOARVLAVERYLGDQQLLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNNTMME 578  
 Db 571 QLOARVLAVERYLGDQQLLGIGWCSGKLICTTAVPMNASWSNKSIDRIIMNNTMME 630  
 Qy 579 IDNTSEIYTLIEESQNOEKNEQELBLDKASIMNFDITKMLY 625

Db 631 IDNTNLIYSLIEDSQIOEENKELLEDKMASIMMFINITNLMWY 677

## RESULT 10

533985  
env polyprotein - human immunodeficiency virus type 1  
C/Species: human immunodeficiency virus type 1, HIV-1  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999  
C/Accession: S33985  
R/Carlini, F.  
Submitted to the EMBL Data Library, November 1991  
A/Reference number: S33979  
A/Accession: S33985  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-851 <CAR>  
A/Cross-references: EMBL:Z11530, NID:660192, PIDN:CAA77628.1, PID:660199  
C/Superfamily: type E retrovirus env polyprotein

Query Match Best Local Similarity 85.5%; Score 2866.5; DB 2; Length 851;  
Matches 543; Conservative 33; Mismatches 46; Indels 25; Gaps 5;

Qy 2 EKLMTVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 61  
Db 32 EKLMTVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 91  
Qy 62 HENMKNNNNVEQOEIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG--TWER 119  
Db 92 HENMKNNNNVEQOEIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG--TWER 151  
Qy 120 GEINCSFNITTSIRDEYQKEVALFYKLDVYXINNNTSYRLISCDTSVITQACPKISFE 179  
Db 152 GEINCSFNITTSIRDEYQKEVALFYKLDVYXINNNTSYRLISCDTSVITQACPKISFE 211  
Qy 180 PIPHYCAPAGFALLKCNKDTFNGKXPCKNVSTXOCTHGIRPVSTOLLNGLSIAEEBV 239  
Db 212 PIPHYCAPAGFALLKCNKDTFNGKXPCKNVSTXOCTHGIRPVSTOLLNGLSIAEEBV 271  
Qy 240 IRSNFINNAKTIIVOLKESVEINCTRPNN-----GAGIRQAH 280  
Db 272 IRSNFINNAKTIIVOLKESVEINCTRPNN-----GAGIRQAH 331  
Qy 281 NISPAKNNNDTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTOLF 339  
Db 332 NISPAKNNNDTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTOLF 391  
Qy 340 NSTWNNNTSGSNTEG--NTITLPCRIRKQIIMMOEGVKAMYPPIRGQIRCSSNITGLL 398  
Db 392 NSTW--STKGSNNTSGSDITLPCRIRKQIIMMOEGVKAMYPPIRGQIRCSSNITGLL 449  
Qy 399 TRDGINENGTEIRPPGGMDRDMWRBSELYKYVKLEPLGVAATPKCRARVVOEKAVG 458  
Db 450 TRDGSNNSEIRPPGGMDRDMWRBSELYKYVKLEPLGVAATPKCRARVVOEKAVG 509  
Qy 459 IGAFLFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQTLWGK 518  
Db 510 IGAFLFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQTLWGK 569  
Qy 519 QLOARVLAVERYLDOQLLGIGWCSGKLICTTAVPWNASNSKSLDIRIMNMTWEMERE 578  
Db 570 QLOARVLAVERYLDOQLLGIGWCSGKLICTTAVPWNASNSKSLDIRIMNMTWEMERE 629  
Qy 579 IDNTSBIYTLIESQNOEKNEOELLELDKMASIMMFINITNLMWY 625  
Db 630 IDNTSBIYTLIESQNOEKNEOELLELDKMASIMMFINITNLMWY 676

RESULT 11  
VCLJ3M  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)  
N/Contains: coat protein gp120; coat protein gp41

C/Species: human immunodeficiency virus type 1, HIV-1  
C/Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 07-Nov-1997  
C/Accession: A24774  
R/Starchich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.  
Cell 45; 637-648, 1986  
A/Title: Identification and characterization of conserved and variable regions in the en  
A/Reference number: A24774; MUID:86218077; PMID:2423250  
A/Accession: A24774  
A/Molecule type: DNA  
A/Residues: 1-856 <SRA>  
A/Cross-references: GB:K03455; GB:M38432; NID:q1906382  
C/Genetics:  
A/Gene: env  
C/Superfamily: type E retrovirus env polyprotein  
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F.1-29/Domain: signal sequence #status predicted <SIG>  
F.30-501/Product: coat protein gp120 #status predicted <GP1>  
F.502-847/Product: coat protein gp41 #status predicted <GP2>  
F.87.134,140,151,155,183,197,234,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match Best Local Similarity 85.4%; Score 2861.5; DB 1; Length 856;  
Matches 547; Conservative 29; Mismatches 47; Indels 31; Gaps 6;

Qy 1 VEKLMVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60  
Db 30 VEKLMVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 89  
Qy 61 EHFMMKNNNNVEQOEIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEGIMERG 120  
Db 90 EHFMMKNNNNVEQOEIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEGIMERG 147  
Qy 121 EIKKCSFNITTSIRDEYQKEVALFYKLDVYXINNNT--TSYRLISCDTSVITQACPK 175  
Db 148 EIKKCSFNITTSIRDEYQKEVALFYKLDVYXINNNT--TSYRLISCDTSVITQACPK 207  
Qy 176 ISFEPPIHYCAPAGFALLKCNKDTFNGKXPCKNVSTXOCTHGIRPVSTOLLNGLSIAE 235  
Db 208 ISFEPPIHYCAPAGFALLKCNKDTFNGKXPCKNVSTXOCTHGIRPVSTOLLNGLSIAE 267  
Qy 236 BEVYIRSDNTNNAKTIIVOLKESVEINCTRPNN-----AGIRQ 277  
Db 268 BEVYIRSDNTNNAKTIIVOLKESVEINCTRPNN-----AGIRQ 327  
Qy 278 AHCHISAKNNNDTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTOLF 337  
Db 328 AHCHISAKNNNDTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTOLF 387  
Qy 338 LFNSTWN--NTEGSNTE--GNTITLPCRIRKQIIMMOEGVKAMYPPIRGQIRCSSN 392  
Db 388 LFNSTWNVTGISTEGNNTENGDTITLPCRIRKQIIMMOEGVKAMYPPIRGQIRCSSN 447  
Qy 393 ITGILLTRDGINENGTEIRPPGGMDRDMWRBSELYKYVKLEPLGVAATPKCRARVVO 452  
Db 448 ITGILLTRDGSNNSEIRPPGGMDRDMWRBSELYKYVKLEPLGVAATPKCRARVVO 507  
Qy 453 EKRAVVG-IGAVFLGFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQ 511  
Db 508 EKRAVVG-IGAVFLGFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQ 567  
Qy 512 LTVWGIQLOARVLAVERYLDOQLLGIGWCSGKLICTTAVPWNASNSKSLDIRIMNMT 571  
Db 568 LTVWGIQLOARVLAVERYLDOQLLGIGWCSGKLICTTAVPWNASNSKSLDIRIMNMT 627  
Qy 572 WMEWEREIDNTSBIYTLIESQNOEKNEOELLELDKMASIMMFINITNLMWY 625  
Db 628 WMEWEREIDNTSBIYTLIESQNOEKNEOELLELDKMASIMMFINITNLMWY 681

RESULT 12  
VCLJ3M  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)  
N/Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C:Accession: A28922  
R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta  
virology 164, 531-536, 1988  
A>Title: Envelope sequences of two new United States HIV-1 isolates.  
A:Reference number: A28922; MUID:88219542; PMID:3369091  
A:Accession: A28922  
A:Molecule type: DNA  
A:Residues: 1-859 <GUR>  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote  
F.1-29/Domain: signal sequence #status predicted <SIG>  
F.30-859/Product: env polyprotein #status predicted <EP>  
F:87,129,135,140,141,146,161,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401

Query Match 85.3%; Score 2859.5; DB 1; Length 859;

Best Local Similarity 83.1%; Pred. No. 9.7e-205;

Matches 545; Conservative 32; Mismatches 46; Indels 33; Gaps 7;

```
QY 2 EKLATVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATACVPTDPNPOEVLNATE 61
DB 31 EKLWTVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATACVPTDPNPOEVLNATE 90
QY 62 HFNMMKNMVEOMEDDIIISLDOSLPCVKLTPLCVTLNCKDV-----NATNTTNDSE 113
DB 91 NFNMMKNMVEOMEDDIIISLDOSLPCVKLTPLCVTLNCKDV-----NATNTTNDSE 150
QY 114 TMEERGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 173
DB 151 EGTIGEMKNCSPNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 210
QY 174 ISFEPPIPIHYCAPAGFAILKNDKTFNGKAPCKNVSTVQCTHGRIPVSTOLLNGSL 233
DB 211 PKISFEPPIHYCAPAGFAILKNDKTFNGKAPCKNVSTVQCTHGRIPVSTOLLNGSL 270
QY 234 ABEVVIRSDNFTNNAKTIIVOLKESVEINCTRPN-----NGAGDI 275
DB 271 ABEVVIRSDNFTNNAKTIIVOLKESVEINCTRPN-----NGAGDI 330
QY 276 ROAHGNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEIYMHSGGGEFFYCN 335
DB 331 ROAHGNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEIYMHSGGGEFFYCN 390
QY 336 TOLFNSTNNNTES--NTEGN-TTLPCKRIKQIIMMOEYGVKAMYPPIRGQIRCS 391
DB 391 TOLFNSTNNNTES--NTEGN-TTLPCKRIKQIIMMOEYGVKAMYPPIRGQIRCS 450
QY 392 NITGLLTRDGIENG-----TEIFRPGGDMDMNMSSELYKRVVYIEPLGVAFTCK 446
DB 441 NITGLLTRDGIENG-----TEIFRPGGDMDMNMSSELYKRVVYIEPLGVAFTCK 510
QY 447 RRVVOREKRAVG-IGAVFLGFGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 505
DB 507 RRVVOREKRAVG-IGAVFLGFGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 566
QY 506 QOQMLQTLTWGIKOLQARVLAVERYLGDQOLLGIGWCSGKLICTTAVPNASWSNKS 565
DB 567 QOQMLQTLTWGIKOLQARVLAVERYLGDQOLLGIGWCSGKLICTTAVPNASWSNKS 626
QY 566 IANNNTWMEEREIDNTYSEIYTLIEESONOEKKEOELLELDKVASLMMNFDITKWL 625
DB 627 IANNNTWMEEREIDNTYSEIYTLIEESONOEKKEOELLELDKVASLMMNFDITKWL 686
```

## RESULT 13

VCLJ3C

env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C:Accession: B28922  
R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta  
virology 164, 531-536, 1988  
A>Title: Envelope sequences of two new United States HIV-1 isolates.  
A:Reference number: A28922; MUID:88219542; PMID:3369091  
A:Accession: B28922  
A:Molecule type: DNA  
A:Residues: 1-861 <GUR>  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote  
F.1-29/Domain: signal sequence #status predicted <SIG>  
F.30-861/Product: env polyprotein #status predicted <EP>  
F:87,129,135,140,143,159,163,187,198,224,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 85.1%; Score 2853; DB 1; Length 861;

Best Local Similarity 82.9%; Pred. No. 3e-204;

Matches 547; Conservative 32; Mismatches 41; Indels 40; Gaps 9;

```
QY 2 EKLATVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATACVPTDPNPOEVLNATE 61
DB 31 EKLWTVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATACVPTDPNPOEVLNATE 90
QY 62 HFNMMKNMVEOMEDDIIISLDOSLPCVKLTPLCVTLNCKDV-----NATNTTNDSE 115
DB 91 NFNMMKNMVEOMEDDIIISLDOSLPCVKLTPLCVTLNCKDV-----NATNTTNDSE 150
QY 116 TMEERGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 175
DB 151 EGTIGEMKNCSPNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 208
QY 176 ISFEPPIPIHYCAPAGFAILKNDKTFNGKAPCKNVSTVQCTHGRIPVSTOLLNGSL 235
DB 209 VSFEPPIPIHYCA-RMFALINCNKKFNCTGVSTVQCTHGRIPVSTOLLNGSL 267
QY 236 BEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNGA-----GDI 275
DB 268 BEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNGA-----GDI 327
QY 276 ROAHGNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEIYMHSGGGEFFYCN 335
DB 328 ROAHGNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEIYMHSGGGEFFYCN 387
QY 336 TOLFNSTNNNTES--NTEGN-TTLPCKRIKQIIMMOEYGVKAMYPPIRGQIRCS 391
DB 388 TOLFNSTNNNTES--NTEGN-TTLPCKRIKQIIMMOEYGVKAMYPPIRGQIRCS 446
QY 392 NITGLLTRDGIENG-----TEIFRPGGDMDMNMSSELYKRVVYIEPLGVAFTCK 446
DB 447 NITGLLTRDGIENG-----TEIFRPGGDMDMNMSSELYKRVVYIEPLGVAFTCK 506
QY 447 RRVVOREKRAVG-IGAVFLGFGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 505
DB 507 RRVVOREKRAVG-IGAVFLGFGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 566
QY 506 QOQMLQTLTWGIKOLQARVLAVERYLGDQOLLGIGWCSGKLICTTAVPNASWSNKS 565
DB 567 QOQMLQTLTWGIKOLQARVLAVERYLGDQOLLGIGWCSGKLICTTAVPNASWSNKS 626
QY 566 IANNNTWMEEREIDNTYSEIYTLIEESONOEKKEOELLELDKVASLMMNFDITKWL 625
DB 627 IANNNTWMEEREIDNTYSEIYTLIEESONOEKKEOELLELDKVASLMMNFDITKWL 686
```

## RESULT 14

VCLJ2

env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C/Date: 17-May-1995 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
 C/Accession: A03976  
 C/Species: human immunodeficiency virus type 1, HIV-1  
 C/Date: 30-Jun-1998 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
 C/Accession: C25523  
 R/Debar, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Strinivasan, A.; Andersen, P.R.; Devare, S.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
 A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human  
 A/Reference number: A94136; MUID:87041461; PMID:3490666  
 A/Accession: C25523  
 A/Molecule type: DNA  
 A/Residues: 1-868 <DES>  
 A/Cross-references: GB:M13137; NID:g326460; PIDN:AAA4311.1; PID:g326467  
 C/Genetics:  
 A/Gene: env  
 C/Superfamily: type B retrovirus env polyprotein  
 C/Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
 F:1-30/Domain: signal sequence #status predicted <SIG>  
 F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>  
 F:510-855/Product: transmembrane glycoprotein #status predicted <TM>  
 F:87,129,140,158,184,190,204,265,292,298,304,334,341,358,364,388,394,400,408,445,458  
 F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.6%; Score 2836; DB 1; Length 855;

Best Local Similarity 83.2%; Pred. No. 5.4e-203; Indels 32; Gaps 9;

Matches 543; Conservative 30; Mismatches 48; Indels 32; Gaps 9;

QY 2 EKLWTVYVGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTHEF 61  
 DB 31 EKLWTVYVGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTHEF 90  
 QY 62 HNNWNNWVQMOEDIIISLWQSLKPCVKLTPLCYTLNCKDVNATNTTNDSEGMT- 120  
 DB 91 HNNWNNWVQMOEDIIISLWQSLKPCVKLTPLCYTLNCKDVNATNTTNDSEGMT- 150  
 QY 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDN-----NNTSYRLISCDTSVITQACPK 175  
 DB 151 EIKKCSFNITTSIRDEVOKEVALFYKLDVXIDN-----NNTSYRLISCDTSVITQACPK 210  
 QY 176 ISFEPPIHYCAPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNGSLAE 235  
 DB 211 VSFEPPIHYCTPGAFALCKNNKTFNKGCTNVSTVQCTHGIRPVVSTOLLNGSLAE 270  
 QY 236 EEVYIRSNFTNNAKTTIIYQKESVEINCTRPNNNG-----AGDIRQ 277  
 DB 271 EEVYIRSNFTNNAKTTIIYQKESVEINCTRPNNNG-----AGDIRQ 330  
 QY 278 AHCHISRAKNDTLKOIYIKLRQF--ENKTI VFNHSSGDEIYVHSHFNCGEFFYC 336  
 DB 331 AHCHISRAKNDTLKOIYIKLRQF--ENKTI VFNHSSGDEIYVHSHFNCGEFFYC 390  
 QY 337 QLFNSTVNN--NTEGSNNTEGS--TTLPCKRIKOIINMVEGKAMVAPPPIRGQIRCSNIT 394  
 DB 391 QLFNSTVNN--NTEGSNNTEGS--TTLPCKRIKOIINMVEGKAMVAPPPIRGQIRCSNIT 447  
 QY 395 GLILTRDGIN--ENGTEIFRPGGDMRDNMRSELYKRVVKKIEPLGVAFTCKRRVORE 453  
 DB 448 GLILTRDGIN--ENGTEIFRPGGDMRDNMRSELYKRVVKKIEPLGVAFTCKRRVORE 507  
 QY 454 KRAVGI--GAVPLGFGAAGSTMGASMTLTVOARLLSGIYQOQNNLLRAIEAOQMLQ 512  
 DB 508 KRAVGI--GAVPLGFGAAGSTMGASMTLTVOARLLSGIYQOQNNLLRAIEAOQMLQ 567  
 QY 513 TVWGIKQIQAARVLAVERYLQDQLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTW 572  
 DB 568 TVWGIKQIQAARVLAVERYLQDQLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTW 627  
 QY 573 MEMEREIDNTYSEIYTLIESQNOQEKQEBLELDKASIMNFDITKMLY 625  
 DB 628 MEMEREIDNTYSEIYTLIESQNOQEKQEBLELDKASIMNFDITKMLY 680

RESULT 15

VCLJH4  
 env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)  
 N/Alternate names: coat polyprotein

N/Contains: coat protein gp120; coat protein gp41  
 C/Species: human immunodeficiency virus type 1, HIV-1  
 C/Date: 30-Jun-1998 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
 C/Accession: C25523  
 R/Debar, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Strinivasan, A.; Andersen, P.R.; Devare, S.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
 A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human  
 A/Reference number: A94136; MUID:87041461; PMID:3490666  
 A/Accession: C25523  
 A/Molecule type: DNA  
 A/Residues: 1-868 <DES>  
 A/Cross-references: GB:M13137; NID:g326460; PIDN:AAA4311.1; PID:g326467  
 C/Genetics:  
 A/Gene: env  
 C/Superfamily: type B retrovirus env polyprotein  
 C/Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
 F:1-521/Product: coat protein gp120 #status predicted <GP1>  
 F:522-868/Product: coat protein gp41 #status predicted <GP2>  
 F:89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,

Query Match 83.4%; Score 2793.5; DB 1; Length 868;

Best Local Similarity 80.1%; Pred. No. 8.1e-200; Indels 37; Gaps 5;

Matches 528; Conservative 36; Mismatches 58; Indels 37; Gaps 5;

QY 4 LWTYVYVGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTHEF 63  
 DB 35 LWTYVYVGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTHEF 94  
 QY 64 HNNWNNWVQMOEDIIISLWQSLKPCVKLTPLCYTLNCKDVNATNTTNDSEGMT- 117  
 DB 95 HNNWNNWVQMOEDIIISLWQSLKPCVKLTPLCYTLNCKDVNATNTTNDSEGMT- 154  
 QY 118 -ERGEIKNCSFNITTSIRDEVOKEVALFYKLDVXIDN-----NNTSYRLISCDTSVITQ 171  
 DB 155 -ERGEIKNCSFNITTSIRDEVOKEVALFYKLDVXIDN-----NNTSYRLISCDTSVITQ 214  
 QY 172 ACPKISFEPPIHYCAPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNG 231  
 DB 215 ACPKISFEPPIHYCTPGAFALCKNNKTFNKGCTNVSTVQCTHGIRPVVSTOLLNG 274  
 QY 232 SLAEEVYIRSNFTNNAKTTIIYQKESVEINCTRPNNNG-----AG 273  
 DB 275 SLAEEVYIRSNFTNNAKTTIIYQKESVEINCTRPNNNG-----AG 334  
 QY 274 DIRAHCHISRAKNDTLKOIYIKLRQF--ENKTI VFNHSSGDEIYVHSHFNCGEFFYC 333  
 DB 335 DIRAHCHISRAKNDTLKOIYIKLRQF--ENKTI VFNHSSGDEIYVHSHFNCGEFFYC 394  
 QY 334 NSTQLFNSTVNNNTEGS-----NTEGSNTTLPCKRIKOIINMVEGKAMVAPPPIRGQ 387  
 DB 395 NSTQLFNSTVNNNTEGS-----NTEGSNTTLPCKRIKOIINMVEGKAMVAPPPIRGQ 454  
 QY 398 RCGSNTTGLILTRDGIN--ENGTEIFRPGGDMRDNMRSELYKRVVKKIEPLGVAFTCKR 447  
 DB 455 RCGSNTTGLILTRDGIN--ENGTEIFRPGGDMRDNMRSELYKRVVKKIEPLGVAFTCKR 514  
 QY 448 RVVOREKRAVG--GAVPLGFGAAGSTMGASMTLTVOARLLSGIYQOQNNLLRAIEAO 506  
 DB 515 RVVOREKRAVG--GAVPLGFGAAGSTMGASMTLTVOARLLSGIYQOQNNLLRAIEAO 574  
 QY 507 QRMQLTVWGIKQIQAARVLAVERYLQDQLLGIWCGSKLICCTAVPWNASWSNKSIDRI 566  
 DB 575 QRMQLTVWGIKQIQAARVLAVERYLQDQLLGIWCGSKLICCTAVPWNASWSNKSIDRI 634  
 QY 567 WNNMTWMEMEREIDNTYSEIYTLIESQNOQEKQEBLELDKASIMNFDITKMLY 625  
 DB 635 WNNMTWMEMEREIDNTYSEIYTLIESQNOQEKQEBLELDKASIMNFDITKMLY 693

Search completed: December 12, 2003, 12:33:34  
 Job time : 15.5506 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:26:14 ; Search time 32.1467 Seconds  
(without alignments)  
5017.086 Million cell updates/sec

Title: US-10-032-162-17  
Perfect score: 3351  
Sequence: 1 VEKLMTVYGVVPMVKRAT.....ELDKWASIMNWFDTKMLMY 625

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
1 number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	98.3	847	15	Q75760 human immun
2	2993.5	89.3	848	15	Q74999 human immun
3	2944.5	87.9	852	15	Q92761 human immun
4	2944.5	87.9	854	15	Q78225 human immun
5	2939.5	87.7	846	15	Q9PXW7 human immun
6	2938.5	87.7	855	15	003805 human immun
7	2936.5	87.6	854	15	040222 human immun
8	2935.5	87.6	838	15	003806 human immun
9	2931	87.4	854	15	Q77368 human immun
10	2929	87.4	854	15	Q92762 human immun
11	2928	87.3	853	15	Q77989 human immun
12	2926	87.3	853	15	003811 human immun
13	2924	87.3	851	15	056110 human immun
14	2917	87.0	847	15	Q8AR21 human immun
15	2911	86.9	849	15	Q8Q851 human immun
16	2906.5	86.7	860	15	Q9YP50 human immun

17	2903.5	86.6	868	15	Q9WJW5 human immun
18	2900.5	86.6	852	15	Q41883 human immun
19	2900	86.5	861	15	Q9WJW5 human immun
20	2899	86.5	855	15	Q8U664 human immun
21	2899	86.5	859	15	Q8Q850 human immun
22	2898	86.5	873	15	Q69994 human immun
23	2896	86.4	843	15	Q70150 human immun
24	2896	86.4	851	15	Q56562 human immun
25	2895	86.4	752	15	Q70604 human immun
26	2895	86.4	752	15	Q70605 human immun
27	2895	86.4	861	15	Q9E527 human immun
28	2894.5	86.4	856	15	Q71270 human immun
29	2894	86.4	861	15	Q9E523 human immun
30	2892	86.3	853	15	Q56108 human immun
31	2892	86.3	863	15	Q42031 human immun
32	2889	86.2	854	15	Q72502 human immun
33	2889	86.2	863	15	Q9WJW8 human immun
34	2887.5	86.2	850	15	Q74812 human immun
35	2887.5	86.2	856	15	Q72993 human immun
36	2886	86.1	854	15	Q85582 human immun
37	2885.5	86.1	840	15	Q802X4 human immun
38	2885.5	86.1	850	15	Q56113 human immun
39	2884.5	86.1	860	15	Q9YP43 human immun
40	2884.5	86.1	860	15	Q9DH01 human immun
41	2884	86.1	851	15	Q8Q852 human immun
42	2883.5	86.0	843	15	Q70008 human immun
43	2881	86.0	856	15	Q92877 human immun
44	2881	86.0	863	15	Q9WJW4 human immun
45	2880.5	86.0	843	15	Q9QKX8 human immun

#### ALIGNMENTS

RESULT 1

ID	Q75760	PRELIMINARY;	PRT;	847 AA.
AC	Q75760;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	Envelope glycoprotein.			
ENV.				
OS	Human immunodeficiency virus 1.			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11676;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=JREF;			
RX	MEDLINE=91043044; PubMed=2172833;			
RA	O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Koyanagi Y., Miles S., Mitsuyasu R.T., Merrill J.E., Vinters H.V., Chen I.S.;			
RT	"Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms.";			
RT	Science 236:819-822(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=JREF;			
RX	MEDLINE=91043044; PubMed=2172833;			
RA	O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Zack J.A., Chen I.S.;			
RT	"HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain.";			
RT	Nature 348:69-73(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=JREF;			
RX	MEDLINE=92092169; PubMed=1684385;			
RA	Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;			
RT	"HIV-1 env sequence variation in brain tissue of patients with AIDS-related neurologic disease.";			
RT	J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).			
RN	[4]			



RP SEQUENCE FROM N.A.  
 RC STRAIN=JREFL;  
 RA Pang S., Vintere H.V., Akashi T., O'Brien W.A., Chen I.S.,  
 RL SubMITTED (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U63632; AAB05604.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 847 AA; 96160 MW; 022D5F24B04FB29P CRC64;

Query Match 98.3%; Score 3293; DB 15; Length 847;  
 Best Local Similarity 96.1%; Pred. No. 9.1e-261;  
 Matches 618; Conservative 0; Mismatches 7; Indels 18; Gaps 1;

QY 1 VEKLMVTVVYGVVPMKEATTTTLCASDAKAYDEVHNVMAATHACVPTDPNPOEVLNVT 60  
 30 VEKLMVTVVYGVVPMKEATTTTLCASDAKAYDEVHNVMAATHACVPTDPNPOEVLNVT 89  
 QY 61 EHFMMKNNMVEQOEEDIIISLMDOSLKPCKVLTPLCVTLNCKVNAATNTTDSSEGMERG 120  
 90 EHFMMKNNMVEQOEEDIIISLMDOSLKPCKVLTPLCVTLNCKVNAATNTTDSSEGMERG 149  
 QY 121 EIKKCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 180  
 150 EIKKCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 209  
 QY 181 IPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXQCTHGIRPVSTOLLNGLAEBEVY 240  
 210 IPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXQCTHGIRPVSTOLLNGLAEBEVY 269  
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 282  
 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 329  
 QY 283 SRAKMNITLKOIVKLRQEFENKTIIVFNHSGGDEIIVHNSGGEFFYCNSTOLFNS 342  
 330 SRAKMNITLKOIVKLRQEFENKTIIVFNHSGGDEIIVHNSGGEFFYCNSTOLFNS 389  
 QY 343 WNNTEGSNNTEGNTITLPCRIKOIINMDEVGKAMAPPIRGDIRGSSNITGLLTR 402  
 390 WNNTEGSNNTEGNTITLPCRIKOIINMDEVGKAMAPPIRGDIRGSSNITGLLTR 449  
 QY 403 GINNGTETIIFPPGGDMRDNRSSELYKKYKVIIEPLGVAFTKCRRVVQREKRAVGIG 462  
 450 GINNGTETIIFPPGGDMRDNRSSELYKKYKVIIEPLGVAFTKCRRVVQREKRAVGIG 509  
 QY 463 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOORMLQLTVMGIKOLA 522  
 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOORMLQLTVMGIKOLA 569  
 QY 523 RVLAVERYLGDQQLIGWCSGKLICTTAVPMNASWSNKSIDRIINNNNTWMEWEEREIDNY 582  
 570 RVLAVERYLGDQQLIGWCSGKLICTTAVPMNASWSNKSIDRIINNNNTWMEWEEREIDNY 629  
 QY 583 TSEIYITLIESQNOOEKNEQELBLDKVASIMNFDITKMLMY 625  
 630 TSEIYITLIESQNOOEKNEQELBLDKVASIMNFDITKMLMY 672

RESULT 2  
 074999 PRELIMINARY; PRT; 848 AA.  
 AC 074999;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE Envelope glycoprotein gp160.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 VC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxId=11676;  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JR-CSF;  
 RX MEDLINE=97063128; PubMed=8906996;  
 RA Klasse P.J., Boyd M.T., Weiss R.A., Schulz T.F.;  
 RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing  
 RT variant of HIV type 1 JR-CSF that infects a range of T cell lines";  
 RL AIDS Res. Hum. Retroviruses 12:347-350(1996).  
 DR EMBL; U45960; AAB05921.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 848 AA; 96433 MW; B240019C8737BCB3 CRC64;

Query Match 89.3%; Score 2993.5; DB 15; Length 848;  
 Best Local Similarity 87.8%; Pred. No. 3.1e-236;  
 Matches 566; Conservative 26; Mismatches 32; Indels 21; Gaps 4;

QY 1 VEKLMVTVVYGVVPMKEATTTTLCASDAKAYDEVHNVMAATHACVPTDPNPOEVLNVT 60  
 30 VEKLMVTVVYGVVPMKEATTTTLCASDAKAYDEVHNVMAATHACVPTDPNPOEVLNVT 89  
 QY 61 EHFMMKNNMVEQOEEDIIISLMDOSLKPCKVLTPLCVTLNCKVNAATNTTDSSEGMERG 120  
 90 EHFMMKNNMVEQOEEDIIISLMDOSLKPCKVLTPLCVTLNCKVNAATNTTDSSEGMERG 149  
 QY 121 EIKKCSFNITTSIRDEVOKEYALFYKLDVYXIDNNTSYRLISGDTSVITQACPKISFE 179  
 150 EIKKCSFNITTSIRDEVOKEYALFYKLDVYXIDNNTSYRLISGDTSVITQACPKISFE 209  
 QY 181 IPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXQCTHGIRPVSTOLLNGLAEBEVY 240  
 210 IPIHYCAPAGFAILKCNKDTFNGKPKCKNVSTXQCTHGIRPVSTOLLNGLAEBEVY 269  
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281  
 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 329  
 QY 283 SRAKMNITLKOIVKLRQEFENKTIIVFNHSGGDEIIVHNSGGEFFYCNSTOLFNS 341  
 330 SRAKMNITLKOIVKLRQEFENKTIIVFNHSGGDEIIVHNSGGEFFYCNSTOLFNS 389  
 QY 343 WNNTEGSNNTEGNTITLPCRIKOIINMDEVGKAMAPPIRGDIRGSSNITGLLTR 400  
 390 WNNTEGSNNTEGNTITLPCRIKOIINMDEVGKAMAPPIRGDIRGSSNITGLLTR 448  
 QY 403 GINNGTETIIFPPGGDMRDNRSSELYKKYKVIIEPLGVAFTKCRRVVQREKRAVGIG 460  
 450 GINNGTETIIFPPGGDMRDNRSSELYKKYKVIIEPLGVAFTKCRRVVQREKRAVGIG 508  
 QY 463 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOORMLQLTVMGIKOLA 520  
 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOORMLQLTVMGIKOLA 568  
 QY 523 RVLAVERYLGDQQLIGWCSGKLICTTAVPMNASWSNKSIDRIINNNNTWMEWEEREID 580  
 569 RVLAVERYLGDQQLIGWCSGKLICTTAVPMNASWSNKSIDRIINNNNTWMEWEEREID 628  
 QY 581 TSEIYITLIESQNOOEKNEQELBLDKVASIMNFDITKMLMY 625  
 629 TSEIYITLIESQNOOEKNEQELBLDKVASIMNFDITKMLMY 673



## RESULT 3

092761

ID 092761 PRELIMINARY; PRT: 852 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Envelope glycoprotein.  
 GN ENV.

OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPHMS2.9;

RX MEDLINE=98178716; PubMed=9519894;  
 RA McCutchan F.B., Sanders-Buell E., Salminen M.O., Carr J.K.,  
 RA Sheppard W.H.;

"Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in San Francisco Men's Health Study participants";  
 AIDS Res. Hum. Retroviruses 14:329-337(1998).

DR EMBL, AF025750; AAC40588.1; -

DR InterPro; IPR000328; Env\_GP41.

DR InterPro; IPR000777; GAP\_dhdtogenase.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

DR PROSITE; PS00071; GAPDH; 1.

KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB0611B617 CRC64;

Query Match 87.9%; Score 2944.5; DB 15; Length 852;

Best Local Similarity 86.1%; Pred. No. 3.2e-232; Indels 25; Gaps 6;

Matches 557; Conservative 25; Mismatches 40; Indels 25; Gaps 6;

3 KLMVTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 62

32 QLMVTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 91

63 FNMWKNWVQWQEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 121

92 FNMWKNWVQWQEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 151

122 IKNSFNITTSIRDEVOKEVALFYKLDVYXID--NNNTSYRLISCDTSVITQACPKISFE 179

152 IKNSFNITTSIRDEVOKEVALFYKLDVYXID--NNNTSYRLISCDTSVITQACPKISFE 211

180 PIPHYCAPAGPAILKCNKDTFNGKPCKNVSTXOCTHGIRPVVSTOLLNGSLAESEV 239

212 PIPHYCAPAGPAILKCNKDTFNGKPCKNVSTXOCTHGIRPVVSTOLLNGSLAESEV 271

240 IRSDFNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHCN 281

272 IRSDFNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHCN 331

282 ISRAKNDTLKQIVIKLAEQENKTIIVNHSQGPDEIVMHSFNGGSEFFYCNSTQULNS 341

332 ISRAKNDTLKQIVIKLAEQENKTIIVNHSQGPDEIVMHSFNGGSEFFYCNSTQULNS 391

342 TWNNNTGSSNTEGN--TITLPCRIRKQIINMOWEKGKMYAPPIRGQIRCSNITGLLT 399

392 TW--NDTESSNTRNDTITLPCRIRKQIINMOWEKGKMYAPPIRGQIRCSNITGLLT 450

400 RDGGINENG-TEIFRPGGDMRDNNRSELVYKVKVIEPLGVAPTKCRRVVQREKRAVG 458

451 RDGGINENG-TEIFRPGGDMRDNNRSELVYKVKVIEPLGVAPTKCRRVVQREKRAVG 510

459 IGAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTVWGIK 518

511 IGAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTVWGIK 570

519 QLOARVLAVERYLADQDLIGWCGSKLICCTAVPMNASWSNKSIDRIIMNNTWMEWERE 578

## RESULT 4

078225

ID 078225 PRELIMINARY; PRT: 854 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Env protein.  
 GN ENV.

OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ba-1;

RX MEDLINE=91289160; PubMed=1905842;  
 RA Hwang S.S., Boyle T.J., Lyeily H.K., Cullen B.R.;

"Identification of the envelope V3 loop as the primary determinant of RT cell tropism in HIV-1";  
 Science 253:71-74(1991).

DR EMBL; M63929; AAA75116.1; -

DR InterPro; IPR000328; Env\_GP41.  
 DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.

KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 854 AA; 97186 MW; 005FC8B1B5CDBF8 CRC64;

Query Match 87.9%; Score 2944.5; DB 15; Length 854;

Best Local Similarity 85.5%; Pred. No. 3.2e-232; Indels 25; Gaps 3;

Matches 555; Conservative 25; Mismatches 44; Indels 25; Gaps 3;

2 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 61

31 EKLWTVYVGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 90

62 FNMWKNWVQWQEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 115

91 FNMWKNWVQWQEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGTMERGE 150

116 TMRGEIKNSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSYRLISCDTSVITQACPK 174

151 TMRGEIKNSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSYRLISCDTSVITQACPK 210

175 KISFEPIPIHYCAPAGPAILKCNKDTFNGKPCKNVSTXOCTHGIRPVVSTOLLNGSLA 234

211 KISFEPIPIHYCAPAGPAILKCNKDTFNGKPCKNVSTXOCTHGIRPVVSTOLLNGSLA 270

235 EEEVVISDNTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIR 276

271 EEEVVISDNTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIR 330

277 QAHNINISAKANDTLKQIVIKLAEQENKTIIVNHSQGPDEIVMHSFNGGSEFFYCNST 336

331 QAHNINISAKANDTLKQIVIKLAEQENKTIIVNHSQGPDEIVMHSFNGGSEFFYCNST 390

337 QLFNSTWNNNTGSSNTEGN--TITLPCRIRKQIINMOWEKGKMYAPPIRGQIRCSNITGL 396

391 QLFNSTWNNNTGSSNTEGN--TITLPCRIRKQIINMOWEKGKMYAPPIRGQIRCSNITGL 450

397 LLTDDGNGINENG-TEIFRPGGDMRDNNRSELVYKVKVIEPLGVAPTKCRRVVQREKRA 456

451 LLTDDGNGINENG-TEIFRPGGDMRDNNRSELVYKVKVIEPLGVAPTKCRRVVQREKRA 510

457 VGIAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTVWGIK 516

```

Db      511 VGIAGVLELGFAGASTWGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQHLLQLTWG 570
Qy      517 IKQOAVLAVERLYLGOOQLLGICGSGKILCTTAVPNNASMSKSLDRINNNMTWMEWE 576
Db      571 IKQOAVLAVERLYLGOOQLLGICGSGKILCTTAVPNNASMSKSLDRINNNMTWMEWE 630
Qy      577 REINNTSEIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 625
Db      631 REINNTSEIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 679

```

## RESULT 5

```

Q9PYW7 ID 09PYW7 PRELIMINARY; PRT; 846 AA.
AC 09PYW7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DT GP160, envelope glycoprotein.
OS ENV.
OC Human immunodeficiency virus 1.
OX Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN NCBI_TaxID=11676;
RX MEDLINE=93189881; PubMed=8446773;
RA Schmidtayerova H., Gayet O., Guetart N., Bolmont C., Hirsch I.,
  Chermann J.C.;
RT "Characterization of HIV-1-PAR, a macrophage-tropic strain: cell
  tropism, virus/cell entry and nucleotide sequence of the envelope
  glycoprotein."
RL Res. Virol. 144:21-26(1993).
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR007777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

```

```

Query Match      87.7%; Score 2939.5; DB 15; Length 846;
Best Local Similarity 86.2%; Pred. No. 8.1e-232;
Matches 555; Conservative 30; Mismatches 36; Indels 23; Gaps 6;

Qy      2 EKLWTVYVGVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      31 DKLWTVYVGVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy      62 HFNWKNMNVQOMQEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTNDSEGTMEG 120
Db      91 YFNWKNMNVQOMQEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTNDSEGTMEG 150
Qy      121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNTSYRLISCDTSYTTQACPKISFEP 180
Db      151 EIKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNTSYRLISCDTSYTTQACPKISFEP 210
Qy      181 IPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNLSLEBEVVI 240
Db      211 IPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNLSLEBEVVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHNI 282
Db      271 RSDNFTNNAKTIIVOLKESVEIKCTRPNNTNRKSIPIGGRAPYTTGEIIGDIRQAHCTI 330
Qy      283 SRAKMDTLKOIVIKLRQFENKTIIVNHSAGDPEIVMHSFNGGGEFFYCNSTQLENST 342
Db      331 SKTMTMENFKOIVIKLRQFENKTIIVNHSAGDPEIVMHSFNGGGEFFYCNSTQLENST 390
Qy      343 MNNTEGNTNTEGNTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGLLTRDG 402
Db      391 W-NDTEGNDTE-RITLTPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGLLTRDG 448
Qy      403 G-INENGTEIFRPGGDMRDNRSELVYKVKVIEPLGVAPTCKRRVVOREKRAVGGA 461

```

```

Db      449 GNNNNNGTEIFRPVGGDMRDNRSELVYKVKVIEPLGVAPTCKRRVVOREKRAVGGA 508
Qy      462 VPLGFLGAASMTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTWGIKOLO 521
Db      509 MFLGFLGAASMTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTWGIKOLO 568
Qy      522 ARVLAVERLYLGOOQLLGICGSGKILCTTAVPNNASMSKSLDRINNNMTWMEWEIDN 581
Db      569 ARVLAVERLYLGOOQLLGICGSGKILCTTAVPNNASMSKSLDRINNNMTWMEWEIDN 627
Qy      582 YTSEIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 625
Db      628 YTNLIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 671

```

## RESULT 6

```

Q03805 ID 003805 PRELIMINARY; PRT; 855 AA.
AC 003805;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OX Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN NCBI_TaxID=11676;
RX MEDLINE=93189881; PubMed=8446773;
RA Schmidtayerova H., Gayet O., Guetart N., Bolmont C., Hirsch I.,
  Chermann J.C.;
RT "Characterization of HIV-1-PAR, a macrophage-tropic strain: cell
  tropism, virus/cell entry and nucleotide sequence of the envelope
  glycoprotein."
RL Res. Virol. 144:21-26(1993).
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR007777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 855 AA; 96988 MW; EC631A3B84180C8D CRC64;

```

```

Query Match      87.7%; Score 2938.5; DB 15; Length 855;
Best Local Similarity 85.1%; Pred. No. 9.9e-232;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

Qy      2 EKLWTVYVGVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      31 DKLWTVYVGVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy      62 HFNWKNMNVQOMQEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTNDSEGTMEG 115
Db      91 YFNWKNMNVQOMQEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTNDSEGTMEG 150
Qy      116 TMRGEIKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNN-NTSYRLISCDTSYTTQAC 174
Db      151 MVGGEKKNCSFNITTSIRDEVOKEVALFYKLDVYKIDNN-NTSYRLISCDTSYTTQAC 210
Qy      175 KISFEPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNLSLA 234
Db      211 KVSFEPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNLSLA 270
Qy      235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
Db      271 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTNRKSIPIGGRAPYTTGEIIGDIR 330
Qy      277 QAHNISRKMDTLKOIVIKLRQFENKTIIVNHSAGDPEIVMHSFNGGGEFFYCNST 336
Db      331 QAHNISRKMDTLKOIVIKLRQFENKTIIVNHSAGDPEIVMHSFNGGGEFFYCNST 390
Qy      337 QLFNSTNNTEGNTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGL 396
Db      391 QLFNSTNNTEGNTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGL 450
Qy      397 LLTRDGINENGTEIFRPGGDMRDNRSELVYKVKVIEPLGVAPTCKRRVVOREKRA 456

```

```

Db      451 LITRGGGPEDNKTEVFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREKKA 510
QY      457 VGIQAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLDLTWG 516
Db      511 VGIQAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLDLTWG 570
QY      517 IKQOLARVLAVERYLADQDLGIMGCGSKLICCTAVPNNASVSNKSLDIRNNMTWMEWE 576
Db      571 IKQOLARVLAVERYLADQDLGIMGCGSKLICCTAVPNNASVSNKSLDIRNNMTWMEWE 630
QY      577 REIDNVTSEIYTLIESQNOQEKNOBELLDKMASLNNWPDITKWLWY 625
Db      631 REINNVTIYSLIESQNOQEKNOBELLDKMASLNNWPDITKWLWY 679

RESULT 7
ID      040222 PRELIMINARY; PRT; 854 AA.
Q040222 040222;
01-JAN-1998 (TREMBLrel. 05, Created)
01-JAN-1998 (TREMBLrel. 05, Last sequence update)
01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Env polyprotein.
GN      Env.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxId=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AD8;
RX      MEDLINE=96432129; PubMed=8835195;
RA      Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA      Martin M.A., Peden K.W.;
RT      "Construction and characterization of a stable full-length macrophage-
RT      tropic HIV type 1 molecular clone that directs the production of high
RT      titers of progeny virions."
RL      AIDS Res. Hum. Retroviruses 12:191-194 (1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AD8;
RA      Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA      Martin M.A., Peden K.W.C.;
RL      Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF004394; AAB64170.1; -
DR      InterPro; IPR000328; Env_GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 854 AA; 97291 MW; 06C45E89103C6C12 CRC64;

Query Match      87.6%; Score 2936.5; DB 15; Length 854;
Best Local Similarity 86.2%; Pred. No. 1.4e-231;
Matches 562; Conservative 18; Mismatches 43; Indels 29; Gaps 8;

QY      1 VEKLAVTVYVYGVPMKEATTTLPFCASDAKAYDEVHANNWATACVPTDPNPOEVLEANT 60
Db      30 VENVLVTVYVYGVPMKEATTTLPFCASDAKAYDEVHANNWATACVPTDPNPOEVLEANT 89
QY      61 EHFNNMKNMNVQWOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATTTTDSSEGTMER 119
Db      90 EHFNNMKNMNVQWOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATTTTDSSEGTMER 147
QY      120 GEIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db      148 GEIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 207
QY      180 PPIHYCAPGPAFALIKCNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 239
Db      208 PPIHYCAPGPAFALIKCNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 267
QY      240 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCHN 281

```

```

Db      268 IRSSNFTDNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGCGRAFYTTGDIIGDIRQAHCHN 327
QY      282 IRSRAKNTDLKQIYKLEOF-ENKTIYFNHSSGGDEIWHMSFVCGGEPFCNSTOLFEN 340
Db      338 ISRTKMNNTLQIATKLEQFGNNNTIYFNSSGGDEIWHMSFVCGGEPFCNSTOLFEN 387
QY      341 STWNN-----NTEGSNTEGN-TITLPCRIRKQIINMMQEVGKAMYAPPIRGQIRCSNITG 395
Db      368 STWNNNGTWNLTQSGTSGNDITLPCRIRKQIINMMQEVGKAMYAPPIRGQIRCSNITG 447
QY      396 LILTRDGGINENG-TEIFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREK 454
Db      448 LILTRDGGNNHNDIETFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREK 507
QY      455 RAVG-IGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLDLTWG 513
Db      508 RAVGTIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLDLTWG 567
QY      514 VGIQOLARVLAVERYLADQDLGIMGCGSKLICCTAVPNNASVSNKSLDIRNNMTWMEWE 573
Db      568 VGIQOLARVLAVERYLADQDLGIMGCGSKLICCTAVPNNASVSNKSLDIRNNMTWMEWE 627
QY      574 EMEREIDNVTSEIYTLIESQNOQEKNOBELLDKMASLNNWPDITKWLWY 625
Db      628 EMEREIDNVTGILYTLIESQNOQEKNOBELLDKMASLNNWPDITKWLWY 679

RESULT 8
ID      Q03806 PRELIMINARY; PRT; 838 AA.
Q03806 003806;
01-NOV-1996 (TREMBLrel. 01, Created)
01-NOV-1996 (TREMBLrel. 01, Last sequence update)
01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Env protein.
GN      Env.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxId=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
RA      Beaver B.;
RL      Submitted (JUN-1991) to the EMBL/Genbank/DBJ databases.
DR      EMBL; M68894; AAA44196.1; -
DR      InterPro; IPR000328; Env_GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      SEQUENCE 838 AA; 95329 MW; 11044B1CA10CFB4D CRC64;

Query Match      87.6%; Score 2935.5; DB 15; Length 838;
Best Local Similarity 85.8%; Pred. No. 1.7e-231;
Matches 551; Conservative 24; Mismatches 40; Indels 27; Gaps 2;

QY      2 EKLAVTVYVYGVPMKEATTTLPFCASDAKAYDEVHANNWATACVPTDPNPOEVLEANT 61
Db      31 EKLAVTVYVYGVPMKEATTTLPFCASDAKAYDEVHANNWATACVPTDPNPOEVLEANT 90
QY      62 EHFNNMKNMNVQWOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATTTTDSSEGTMER 121
Db      91 EHFNNMKNMNVQWOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATTTTDSSEGTMER 141
QY      122 IKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEPI 181
Db      142 MKNCSFNITTNIRGVQKEVALFYELDIVIDNNKDSYRLISCDTSVITQACPKISFEPI 201
QY      182 PPIHYCAPGPAFALIKCNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 241
Db      202 PPIHYCAPGPAFALIKCNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 261
QY      242 SDFNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCHN 283

```

```

Db      262  SENFTNNAKTIIVQJNESVEINCTRPNNTRKSIINGPRAFYTTGELIGDIRQAHCNLS 321
Qy      284  RAKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIYVHNSFNGGFEFFCNSTOLFNSTW 343
Db      322  RAKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIYVHNSFNGGFEFFCNSTOLFNSTW 381
Qy      344  NNNTGSSNTEGNTITLPCRIRKQIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 403
Db      382  NVTESNNTVENNTITLPCRIRKQIINNMQEVGKAMVAPPIRGQIRCSNITGLLTRDG 441
Qy      404  INENGTETFRPGGDMRDNMRSELYKYKVIKIEPLGVAPTKCRVVOREKRAVGIGAV 463
Db      442  PEDKTEVFRPGGDMRDNMRSELYKYKVIKIEPLGVAPTKCRVVOREKRAVGIGAV 501
Qy      464  LGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQRLQITWGIKOLQAR 523
Db      502  LGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQRLQITWGIKOLQAR 561
Qy      524  VLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREIDNYT 583
Db      562  VLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREIDNYT 621
Qy      584  SEIYTLIEESQNOQEKNEQELLEDKVASLMMNFDTIKWLY 625
Db      622  SIYSLIEESQNOQEKNEQELLEDKVASLMMNFDTIKWLY 663

```

## RESULT 9

```

Qy      077368  PRELIMINARY; PRT; 849 AA.
ID      077368;
AC      077368;
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT      01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE      Env polypeptide.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Buckler-White A., Theodore T.
RL      Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.
DR      EMBL; M38430; AAA44073.1; -.
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ      SEQUENCE 849 AA; 97166 MW; 8DBEGBA86AA92467 CRC64;

```

```

Query Match      87.5%; Score 2931; DB 15; Length 849;
Best Local Similarity 86.4%; Pred. No. 4e-231;
Matches 557; Conservative 28; Mismatches 38; Indels 22; Gaps 5;

Qy      2  EKLWTVYGVYVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      31  EKLWTVYGVYVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy      62  HNNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSEGTMRG 120
Db      91  NFNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSEGTMRG 150
Qy      121  EIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSYTTQACPKISFE 180
Db      151  EIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSYTTQACPKISFE 210
Qy      181  IPIHYCAPAGFAILLKNDKXTNGKXPKCNVSTXOCTHGRIPVSTOLLNLSLSEEVYI 240
Db      211  IPIHYCAPAGFAILLKNDKXTNGKXPKCNVSTXOCTHGRIPVSTOLLNLSLSEEVYI 270
Qy      241  RSDNFTNNAKTIIVQJNESVEINCTRPN-----NNGAG-----DIRQAHCN 282

```

```

Db      271  RSDNFTNNAKTIIVQJNESVEINCTRPNIYRKGRHIGPRAFHTRQIENIROAHCN 330
Qy      283  SRKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIYVHNSFNGGFEFFCNSTOLFNST 342
Db      331  SRKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIYVHNSFNGGFEFFCNSTOLFNST 390
Qy      343  MN--NTGSSNTEGNTITLPCRIRKQIINNMQEVGKAMVAPPIRGQIRCSNITGLLTR 400
Db      391  MNSTEGKSGNNTNEN--ITLPCRIRKQIINNMQEVGKAMVAPPIRGQIRCSNITGLLTR 449
Qy      401  DGINENGTETFRPGGDMRDNMRSELYKYKVIKIEPLGVAPTKCRVVOREKRAVGIG 460
Db      450  DGINENGTETFRPGGDMRDNMRSELYKYKVIKIEPLGVAPTKCRVVOREKRAVGIG 509
Qy      461  AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQRLQITWGIKOL 520
Db      510  AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQRLQITWGIKOL 569
Qy      521  QARVLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 580
Db      570  QARVLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 629
Qy      581  NYTSEIYTLIEESQNOQEKNEQELLEDKVASLMMNFDTIKWLY 625
Db      630  NYTSEIYTLIEESQNOQEKNEQELLEDKVASLMMNFDTIKWLY 674

```

## RESULT 10

```

Qy      092762  PRELIMINARY; PRT; 854 AA.
ID      092762;
AC      092762;
DT      01-NOV-1998 (TREMblrel. 08, Created)
DT      01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT      01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE      Envelope glycoprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=SEPMH6.3;
RC      MEDLINE=9818716; PubMed=9519894;
RA      McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RT      "Diversity of the human immunodeficiency virus type 1 envelope
RT      glycoprotein in San Francisco Men's Health Study participants.";
RL      AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR      EMBL; AF025754; AAC40591.1; -.
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ      SEQUENCE 854 AA; 97287 MW; AFD8A82B62C5B27D CRC64;

```

```

Query Match      87.4%; Score 2929; DB 15; Length 854;
Best Local Similarity 85.4%; Pred. No. 5.9e-231;
Matches 554; Conservative 30; Mismatches 39; Indels 26; Gaps 6;

Qy      2  EKLWTVYGVYVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      32  EKLWTVYGVYVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 91
Qy      62  HNNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSEGTMR 119
Db      92  NFNMMKNMVEQWQEDIIISLDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSEGTMR 151
Qy      120  GEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSYTTQACPKISFE 179
Db      152  GEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSYTTQACPKISFE 211

```

QY 180 PIPHYCAPAGPAILKCDKDTFNGCKPCKNVSTQCTHGIRPVVSTOLLNGSLAEEVY 239  
 Db 212 PIPHYCAPAGPAILKCDKDKFSGKGPCTINVTQCTHGIRPVVSTOLLNGSLAEEVY 271  
 QY 240 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIOAHCN 281  
 Db 212 IRSDFNTNNAKTIIVOLKEPVEINCTRPNNNRKSIPIGPGAFYTTGELLIGNIOAHCN 331  
 QY 282 ISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIWMHSGGEPFYCNSTOLFNS 341  
 Db 332 ISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIWMHSGGEPFYCNSTOLFNS 391  
 QY 342 TAWN-NTGSSNTEGN-TITLPCRIKOIINMOWEYKAMYPAPIRGOIRCSNITGLLLT 399  
 Db 392 TAWNTESEGVNTEGNDTITLLCRIKOIYVNMWQEVGKAMYPAPIRGOIRCSNITGLLLT 451  
 QY 400 RDGGINCGT---ELFRPGGDMRNMRSSELYKYRVNTEPELVAPTKCKRVVQREKRA 456  
 Db 452 RDGG-NSNGSNKAEVFRPGGDMRNMRSSELYKYRVNTEPELVAPTKCKRVVQREKRA 510  
 QY 457 VGIGAVFGLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRLQITWVG 516  
 Db 511 AGLGVMFGLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRLQITWVG 570  
 QY 517 IKQLOARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPMNASNSKSLDRINNMWTWME 576  
 Db 571 IKQLOARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPMNASNSKSLDRINNMWTWME 630  
 QY 577 REINDYTSIYTLIESONQOEKNEBELDKMASLMMFDTITWMLY 625  
 Db 631 REINDYTSIYTLIESONQOEKNEBELDKMASLMMFDTITWMLY 679

## RESULT 11

Q77989 PRELIMINARY; PRT; 863 AA.  
 AC Q77989;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Envelope glycoprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF128A;  
 MEDLINE=91056585; PubMed=2243391;  
 RT "The viral envelope gene is involved in macrophage tropism of a human immunodeficiency virus type 1 strain isolated from brain tissue."  
 RL J. Virol. 64:6148-6153 (1990).  
 DR EMBL; M95292; AAA44331.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
 KW SEQUENCE 863 AA; 98222 MW; 0157397FAB84C5D7 CRC64;

Query Match 87.4%; Score 2928; DB 15; Length 863;  
 Best Local Similarity 84.4%; Pred. No. 7.2e+231;  
 Matches 557; Conservative 25; Mismatches 40; Indels 38; Gaps 7;

QY 2 EKLWTVYGVVPWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPQEVLENT 61  
 Db 31 EKLWTVYGVVPWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPQEVLENT 90  
 QY 62 HNNMKNNMVEOMODIISIMDOSLKPCVKLTPLCVTLNCD---VNAITNTNSSEGIM 117  
 Db 91 HNNMKNNMVEOMODIISIMDOSLKPCVKLTPLCVTLNCD---VNAITNTNSSEGIM 149

QY 118 -----ERGEINQCFENITTSIRDEVQEVAFYKLDIVXIDNNNTSYRLISCDTSVITQA 172  
 Db 150 SLSRGETEIKNCSFNITTSIRDKQKEYALFYKLDIVIPINNNTTYRLINCNSTITQA 209  
 QY 173 CPKISFEPIPIHYCAPAGPAILKCDKTFNGKAPCKNVSTQCTHGIRPVVSTOLLNGS 232  
 Db 210 CPKISFEPIPIHYCAPAGPAILKCDKCKFNKGPCKNVSTQCTHGIRPVVSTOLLNGS 269  
 QY 223 LAEEVVRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GA--GD 274  
 Db 270 LAEEVVRSDNFTNNAKTIIVOLNESVINCTRPNNNRKSIINGPGAIYTTCAITGD 329  
 QY 275 IROAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIWMHSGGEPFYCN 334  
 Db 330 IROAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIWMHSGGEPFYCN 389  
 QY 335 STOLFNSITWN-----NTEGSSNTEGNTITLPCRIKOIINMOWEYKAMYPAPIRG 385  
 Db 390 TTKLFNSITWNDRLEFNSITWNTEGNTITLPCRIKOIINMOWEYKAMYPAPIRG 448  
 QY 386 QIRCSNITGLLLTRDGINENGTEIFRPGGDMRNMRSSELYKYRVNTEPELVAPTKC 445  
 Db 449 QIRCSNITGLLLTRDGINENGTEIFRPGGDMRNMRSSELYKYRVNTEPELVAPTKC 508  
 QY 446 KRRVQREKRAVIGAVFGLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEA 505  
 Db 509 KRRVQREKRAVIGAVFGLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEA 568  
 QY 506 QORMLQITWVGIKOLOARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPMNASNSKSLDR 565  
 Db 569 QORMLQITWVGIKOLOARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPMNASNSKSLDR 628  
 QY 566 IMNNMTWMEEREIDNYTSIYTLIESONQOEKNEBELDKMASLMMFDTITWMLY 625  
 Db 629 IMNNMTWMEEREIDNYTSIYTLIESONQOEKNEBELDKMASLMMFDTITWMLY 688

## RESULT 12

Q003811 PRELIMINARY; PRT; 853 AA.  
 AC Q003811;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Surface envelope glycoprotein.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ADA;  
 MEDLINE=91195299; PubMed=2014229;  
 RX Westervelt P., Gendelman H.E., Ratner L.;  
 RT "Identification of a determinant within the human immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101 (1991).  
 DR EMBL; M60472; AAA45065.1;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 DR Pfam; PF00517; GP41.1.  
 KW Envelope protein.  
 SQ SEQUENCE 853 AA; 96998 MW; AD2AP21E2B06AD78 CRC64;

Query Match 87.3%; Score 2926; DB 15; Length 853;  
 Best Local Similarity 85.4%; Pred. No. 1e-230;  
 Matches 556; Conservative 25; Mismatches 42; Indels 28; Gaps 7;

QY 1 VEKLMWTVYGVVPWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPQEVLENT 60  
 Db 30 VEKLMWTVYGVVPWKEATTTLLFCASDAKAYDTEVHNWMAHACVPTDPNPQEVLENT 89

```

Qy 6 EHNMMKNMNVBEOMQDIIILMDQSLKPCVLTPLCVJNCKDV-NANTTNDSGEMER 119
Db 90 ENFRMMKNMNVBEOMHDDIILMDQSLKPCVLTPLCVJNCTDLNNTVNNINNSSEGA-R 147
Qy 120 GEIKNSFNITTSIRBEVOKEVALFKLDVVAIDNNNTSYRLISGDSVITQACPXSFE 179
Db 148 GEIKNSFNITTSIRKVKKDYALPRLDVPIIDNDNTSYRLINNTSTIQACPXSFE 207
Qy 180 PIPIHCAPAGAILKNDKTENGKKPCKNVSTXOCTGIRPVSTOLLNGSLAEEV 239
Db 208 PIPIHCTPAGPAILKCKDKKFKNGTGPCQNVSTVOCTGIRPVSTOLLNGSLAEEV 267
Qy 240 IRSDNFPNNAKTIIVOLKESVEINCRPNMNG-----AGDIRAHCH 281
Db 268 IRSNFPDNANKIIVOLKESVEINCRPNMNRKSIHIGCPAFYTTGIIIGDIRAHCH 327
Qy 282 ISRAKNNDTLKQIVILKREOF-ENKTIYVNHSSGADPEIYVHSPFCGGEFFYCSTOLF 340
Db 328 ISHTKNNTLNQIATLKKEQFGNNKTIYVNSGGDEPEIYVHSPFCGGEFFYCSTOLF 387
Qy 341 STWNN-----NTBSNNTGEGN-TTLPFCRIKQIINMDEVKAMVAPIRQIRCSSNITG 395
Db 388 STWNFNGTWNLTQSNSTEGNDITTLPCRKQIINMDEVKAMVAPIRQIRCSSNITG 447
Qy 396 LLITRROGINENTELFRPGGGMWRNMRSELYTKVVKIPELGAAPKARRVQREKR 455
Db 448 LLITRROGINSSSEIFRPGGGMWRNMRSELYTKVVKIPELGAAPKARRVQREKR 507
Qy 456 AVG-IGAVFLGFGAAGSTGAASMLTYQARLLISGIVOOONNLRAIEAQQEMLUTV 514
Db 508 AVGTIGAMFLGFGAAGSTGAASITLTYQARLLISGIVOOONNLRAIEAQQEMLUTV 567
Qy 515 WGIKQIARVLAVERLYGDOULLGIVGSGKLICTTAPVPMNASWNSKLDIRINNMTWME 574
Db 568 WGIKQIARVLAVERLYRDOULLGIVGSGKLICTTAPVPMNASWNSKTLDMIMDMTWME 622
Qy 575 WBERIDNYSIETTLIBESONQOEKREGBLSELDKMASIMNFDITKTLWT 625
Db 628 WBERIDNYSIETTLIBESONQOEKREGBLSELDKMASIMNFDISNLWT 678

```

Query	Score	Length
07.3%	Score 2924;	DB 15;
056110		
ID	056110	PRELIMINARY;
AC	056110	PRT;
DT	01-JUN-1998 (TrEMBLrel. 06, Created)	851 AA.
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
GN	Envelope glycoprotein.	
OS	Human immunodeficiency virus 1.	
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.	
NCBI	TaxID=11676;	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SFMSH.1;	
RA	MEDLINE=98178716; PubMed=9519894;	
RA	McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,	
RA	Sheppard W.H.;	
RT	"Diversity of the human immunodeficiency virus type 1 envelope	
RT	glycoprotein in San Francisco Men's Health Study participants."	
RL	AIDS Res. Hum. Retroviruses 14:329-337(1998).	
DR	EMBL; AF025756; AAC04593.1;	
DR	InterPro; IPR000328; Env_Gp41.	
DR	InterPro; IPR000173; GAP_Thdogenase.	
DR	InterPro; IPR000777; GP120.	
DR	Pfam; PF00516; GP120; 1.	
DR	Pfam; PF00517; GP41; 1.	
DR	PROSITE; PS00071; GAPDH; 1.	
KW	AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.	
SO	SEQUENCE 851 AA; 96516 MW; 1CDAC83EE6464531 CRC64;	

Best Local Similarity 85.6%; Pred. No. 1.5e-230;  
Matches 555; Conservative 27; Mismatches 40; Indels 26; Gaps 5;

QY	2	KJLWVTVYGVWPKWKAFTTLLFCASDPAKYDTEVHNWATACVPTDNPDEVLNANTE	61
Db	31	DKLWVTVYGVWPKWKAFTTLLFCASDPAKYDTEVHNWATACVPTDNPDEVLNANTE	90
QY	62	HFNNKKNMVEQOMEDIILMDQSLKPCVKTLPCLVTLNCKDV-NATVTVDSGCTMWRG	120
Db	91	NFNNKKNMVEQOMEDIILMDQSLKPCVKTLPCLVTLNCKDLNNDTVTSSSGCTMWRG	150
QY	121	EIKKCSNITTSIDPEVQKEVALYXKLDVYXIDNNNTSYRLISDTSVITQACPISPEP	180
Db	151	EIKKCSNITTSIDPKVQKEVALLHKLDVFPID-NDSYRLVSCNTSVITQACPVSPEP	209
QY	181	IPHYCAPAGFAILKCKNDKTFNGKXPCKNVSTXOCTGIRPEVSTOLNGLASLAEENVYI	240
Db	210	IPHYCAPAGFAILKCKNDKFRNGKGPCTNNVSTVQCTHGIRPEVSTQPLNGLSLAEENVYI	269
QY	241	RSNDFPNNAKTIIYQLKESVEINCTRPNNNG-----AGDIRQAHNI	282
Db	270	RSDFPNTNAKTIIYQLNBSVEINCIIRPNNNRKKSINIGPRAFYTTGHEIIDIDIRQAHNL	329
QY	283	SRACKNDTLQIOVKLREOFENKTIIVFHSGSGPEIYMHSPNCGGEFFYCNSTQLFNST	342
Db	330	SEAKMNHLEQIAKLREOFENKTIIVFOSSGSGPEIYMHSPNCGGEFFYCNSTQLFNST	389
QY	343	WN-----NTEGSGNNTEGNTITTLFCRIKQIIMNMOEVGKAMYAPPIRQIRCSSNITGL	397
Db	390	WSVNSTNNDTEGVNNT-GENITLFCRIKQIIMNMOEVGKAMYAPPIRQIRCSSNITGL	448
QY	398	LTRDQGINENGTEIFRPFGGDMRDNMSSELYKYRVKIEPIGVAPYTKCKRYVOREKAV	457
Db	449	LTRDQGSNNMETETFRPFGGDMRDNMSSELYKYRVKIEPIGVAPYTKCKRYVOREKAV	508
QY	458	GIGAVPIGFGLGAAGSTNGAASMTLLTVQARLLLSGIVOOQNNLLAIEAQOMLQJTVWGI	517
Db	509	GIGAVPIGFGLGAAGSTNGAASMTLLTVQARLLLSGIVOOQNNLLAIEAQOMLQJTVWGI	568
QY	518	KQLQARVLAVERYLGDQOLLGIMWCSGKLICTTAVPNNASVSNKSLDRIMNMTWMEWR	577
Db	569	KQLQARVLAVERYLKDQOLLGIMWCSGKLICTTAVPMTKSMNSKSLKXIMNMTWMEWR	628
QY	578	EIDNYTSEITYLIESONQOEKNOELLIEDPKASLWNVPEITIKWLY	625
Db	629	EINNYTSLIYTLIESONQOEKNOELLIEDPKASLWNVPEITIKWLY	676

RESULT 14				
Q8AR21				
ID	Q8AR21	PRELIMINARY;	PRT;	847 AA.
AC	Q8AR21;			
DT	01-MAR-2003	(TREMBLrel. 23, Created)		
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	Envelope glycoprotein.			
GN	ENV.			
OS	Human immunodeficiency virus 1.			
OC	Viruses; Retroid viruses; Retroviridae; Lentiviruses.			
OX	NCBI_TaxID=11676;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SHIVSF162P3.2;			
RL	Gao F.;			
RA	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF536757; AAN05642.1;..			
SO	SEQUENCE 847 AA; 95879 MW; 9CFF863B97B3BB54C CRC64;			
	Query Match	87.0%;	Score 2917;	DB 15; Length 847;
	Best Local Similarity	85.9%;	Pred. No.5,6e-230;	
	Matches 554; Conservative	27;	Mismatches 42;	Indels 22; Gaps 5
QY	1 VEKLVATVYGVPWKEATITLLFCASDAKAYDTEVHNWYATHACVPTDPNFOEVLLENT 60			

```

Db 30 VEKLVATVYGVPAWKEATTTTLFCASDAKAYDTEVHNWATACVPTDPNPOEIVLENTV
QY 61 EHFNNKNNMVEOMOMEDIISLMDOSLKPCVKLTPLCVTLNCDV-NATTTTDSGTHR 119
Db 90 EEFNNKNNMVEOMOMEDIISLMDOSLEPCVKLTPLCVTLNCDV-NATTTTDSGTHR 149
QY 120 GEIKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 150 GEIKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFE 209
QY 180 PIPIHYCAPAGFAIILCKNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 239
Db 210 PIPIHYCAPAGFAIILCKNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 269
QY 240 IRSDFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAHCHN 281
Db 270 IRSEFTDNVKTIIIVOLKESVEINCTRPNNNRKSIPIGPGKAFYATGDIIGDIRAHCHN 329
QY 282 ISRAKNDTLKQIVILKEQFENKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 341
Db 330 ISGEKNNNTLQIVILKEQFENKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 389
QY 342 TWNNTEGSSNTEGNTITLPCRILKQIINMGEVGAAYAPPIRGQIRCSSNITGLLTRD 401
Db 390 TW-NNTIGPNNNG-TITLPCRILKQIINMGEVGAAYAPPIRGQIRCSSNITGLLTRD 447
QY 402 CGIN-ENGTEIFRPGGDMRDNRSELKYKYVVKIEPLGVAPTKRRVVOREKAVGIG 460
Db 448 GREGVNTTEIFRPGGDMRDNRSELKYKYVVKIEPLGVAPTKRRVVOREKAVGIG 507
QY 461 AVFLGFLGAGSTGMAASMTLTVQARLLSGIVOOQNNLLRAIEAQOQMLQLTWGIKQL 520
Db 508 AVFLGFLGAGSTGMAASMTLTVQARLLSGIVOOQNNLLRAIEAQOQMLQLTWGIKQL 567
QY 521 QARVLAVERYLKDOQLLGIWCGSGKLICTTAVPNNASVNSKSLDRIINNMTMWEREID 580
Db 568 QARVLAVERYLKDOQLLGIWCGSGKLICTTAVPNNASVNSKSLDRIINNMTMWEREID 627
QY 581 NTSIYTLTIESONQOEKNEOELLELDKMASLNMWPTITKWLWY 625
Db 628 NTSIYTLTIESONQOEKNEOELLELDKMASLNMWPTITKWLWY 672

RESULT 15
QY 080851 PRELIMINARY: PRT: 849 AA.
Db 080851 01-JUN-2002 (TREMBlrel. 21, Created)
080851 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
080851 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Gp160.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KR3026;
RA Daniels R.S., Patel D., Xiang Z., Zheng N.N., Kang C.;
RT "Evidence for the spread of immune-escape HIV-1 subtype B in the
RT Korean population."
RU Submitted (OCT-2001) to the EMBL/Genbank/DBD databases.
DR EMBL, AJ174427, CAD10139.1, -.
DR InterPro; IPR000328; Env_Gp41.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR InterPro; IPR000777; Gp120.
DR Pfam; PF00516; Gp120_1.
DR Pfam; PF00517; Gp41_1.
DR PROSITE; PS00071; GAPDH_1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 849 AA, 96428 MW, FD660B39D1895C27 CRC64;

```

```

Query Match 86.9%; Score 2911; DB 15; Length 849;
Best Local Similarity 84.7%; Pred. No. 1.7e-229;
Matches 546; Conservative 33; Mismatches 44; Indels 22; Gaps 3;

QY 2 EKLAVTYYGVPAWKEATTTTLFCASDAKAYDTEVHNWATACVPTDPNPOEIVLENTV 61
Db 31 EPLAVTYYGVPAWKEATTTTLFCASDAKAYDTEVHNWATACVPTDPNPOEIVLENTV 90
QY 62 HFNMMKNNMVEOMOMEDIISLMDOSLKPCVKLTPLCVTLNCDVNTNTTDSGTHRGE 121
Db 91 EFNMMKNNMVEOMOMEDIISLMDOSLKPCVKLTPLCVTLNCDVNTNTTDSGTHRGE 150
QY 122 IKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEPI 181
Db 151 IKNSFNITTSIRDEVOKEVALFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEPI 210
QY 182 PIHYCAPAGFAIILCKNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 241
Db 211 PIHYCAPAGFAIILCKNDKTFNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 270
QY 242 SDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAHCHN 283
Db 271 SANFTDNVKTIIIVOLKESVEINCTRPNNNRKSIPIGPGKAFYATGDIIGDIRAHCHN 330
QY 284 RAKNNDTLKQIVILKEQFENKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 343
Db 331 RAKNNDTLKQIVILKEQFENKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 390
QY 344 NNN--TEGSSNTEGNTITLPCRILKQIINMGEVGAAYAPPIRGQIRCSSNITGLLTR 400
Db 391 DNDTDTKETNDTEG-TITLPCRILKQIINMGEVGAAYAPPIRGQIRCSSNITGLLTR 449
QY 401 DGINENGTETFRPGGDMRDNRSELKYKYVVKIEPLGVAPTKRRVVOREKAVGIG 460
Db 450 DGINENGTETFRPGGDMRDNRSELKYKYVVKIEPLGVAPTKRRVVOREKAVGIG 509
QY 461 AVFLGFLGAGSTGMAASMTLTVQARLLSGIVOOQNNLLRAIEAQOQMLQLTWGIKQL 520
Db 510 ALFLGFLGAGSTGMAASMTLTVQARLLSGIVOOQNNLLRAIEAQOQMLQLTWGIKQL 569
QY 521 QARVLAVERYLKDOQLLGIWCGSGKLICTTAVPNNASVNSKSLDRIINNMTMWEREID 580
Db 570 QARVLAVERYLKDOQLLGIWCGSGKLICTTAVPNNASVNSKSLDRIINNMTMWEREID 629
QY 581 NTSIYTLTIESONQOEKNEOELLELDKMASLNMWPTITKWLWY 625
Db 630 NTSIYTLTIESONQOEKNEOELLELDKMASLNMWPTITKWLWY 674

```

Search completed: December 12, 2003, 12:32:40  
 Job time : 33.1467 secs

**THIS PAGE BLANK (US:TO)**